GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 16:35:35 ; Search time 9440.04 Seconds
(without alignments)
10279.455 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgccctacagcaca.....gcaataacattctaaatatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2052	100.0	2052	6	E40087	E40087 Plant promo
	2	2048.8	99.8	2052	6	E40093	E40093 Plant promo
	3	2038	99.3	2056	6	E40091	E40091 Plant promo
	4	2034	99.1	2048	6	E40090	E40090 Plant promo
	5	2032.4	99.0	2048	6	E40089	E40089 Plant promo
	6	241.2	11.8	2042	6	AR076817	AR076817 Sequence
	7	241.2	11.8	2042	6	E15125	E15125 Promoter. 7
	8	196.6	9.6	247	6	AR076816	AR076816 Sequence
	9	196.6	9.6	247	6	E15124	E15124 Promoter. 7
	10	195.6	9.5	246	6	AR146852	AR146852 Sequence
	11	195.6	9.5	246	6	E55065	E55065 Plant promo
c	12	174	8.5	8005	8	DCA18706	Y18706 Daucus caro
	13	162.4	7.9	2831	6	BD188672	BD188672 Promotor
	14	162.4	7.9	2865	6	BD188674	BD188674 Promotor
	15	162.4	7.9	4886	8	DARGCHS2	D16255 Carrot gCHS
	16	119	5.8	140	6	AR146856	AR146856 Sequence
c	17	119	5.8	140	6	AR146857	AR146857 Sequence
	18	119	5.8	140	6	E55069	E55069 Plant promo
c	19	119	5.8	140	6	E55070	E55070 Plant promo
c	20	106.6	5.2	250029	3	AE014839	AE014839 Plasmodiu
c	21	103.6	5.0	172816	9	AC093899	AC093899 Homo sapi
	22	101.8	5.0	258658	3	AE014832	AE014832 Plasmodiu
	23	100.4	4.9	348174	3	CR382399	CR382399 Plasmodiu
	24	99.4	4.8	8056	6	AX599046	AX599046 Sequence
	25	99.2	4.8	205130	2	AC105425	AC105425 Homo sapi
	26	97.6	4.8	176666	2	CR392006	CR392006 Danio rer
c	27	97	4.7	14867	3	AE001398	AE001398 Plasmodiu
	28	96.6	4.7	67970	3	PFMAL1P3	AL031746 Plasmodiu
	29	96.4	4.7	4601	3	DMU11584	U11584 Drosophila
	30	96.4	4.7	19517	3	DMU37541	U37541 Drosophila
	31	95.8	4.7	64394	9	AL928596	AL928596 Human DNA
	32	95.2	4.6	313050	3	PFA929352	AL929352 Plasmodiu
c	33	93.8	4.6	143331	9	AC091214	AC091214 Homo sapi
c	34	93	4.5	349980	6	AX344555	AX344555 Sequence
c	35	92.6	4.5	180629	2	CR388025	CR388025 Danio rer
	36	92.6	4.5	349751	3	PFMAL4P3	AL035476 Plasmodiu
	37	92.4	4.5	261771	2	CR356223	CR356223 Danio rer
	38	91.8	4.5	8056	6	AX598900	AX598900 Sequence
c	39	91.6	4.5	8056	6	AX599046	AX599046 Sequence
	40	91	4.4	21153	8	YSCMTCG13	L36897 Saccharomyc
c	41	91	4.4	118642	9	AC126283	AC126283 Homo sapi
c	42	90.8	4.4	86826	3	PFMAL3P5	AL034556 Plasmodiu
c	43	90.6	4.4	155106	9	AC104069	AC104069 Homo sapi
c	44	90.6	4.4	178670	9	AC104073	AC104073 Homo sapi
c	45	90.2	4.4	105238	9	AC011458	AC011458 Homo sapi

ALIGNMENTS

RESULT 1
E40087
LOCUS E40087 2052 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40087
VERSION E40087.1 GI:18627203
KEYWORDS JP 2000166577-A/1.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 2052)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 1 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000166577-A/1
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC C12N15/00,
PC C12N5/00,(C12N5/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1). .(2052).
FEATURES Location/Qualifiers
source 1. .2052
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 100.0%; Score 2052; DB 6; Length 2052;
Best Local Similarity 100.0%; Pred. No. 3.3e-310;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

```
Qy        61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
Db        61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy       121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
Db       121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy       181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
Db       181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy       241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
Db       241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy       301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
Db       301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy       361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
Db       361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Qy       421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
Db       421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Qy       481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
Db       481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Qy       541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
Db       541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Qy       601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
Db       601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Qy       661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
Db       661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Qy       721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
Db       721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780

Qy       781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
Db       781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840

Qy       841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
Db       841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900

Qy       901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
Db       901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960

Qy       961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
Db       961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020

Qy      1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
Db      1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080

Qy      1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
Db      1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140

Qy      1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
Db      1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200

Qy      1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
Db      1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260

Qy      1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
Db      1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320

Qy      1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
Db      1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380

Qy      1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
Db      1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440

Qy      1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
Db      1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500

Qy      1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
Db      1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560

Qy      1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
Db      1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620

Qy      1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
Db      1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680

Qy      1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
Db      1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740

Qy      1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
Db      1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800

Qy      1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
Db      1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860

Qy      1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATC 1920
Db      1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATC 1920

Qy      1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
Db      1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980

Qy      1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040
Db      1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040

Qy      2041 ATTCTAAATATC 2052
Db      2041 ATTCTAAATATC 2052
```

RESULT 2
E40093
LOCUS E40093 2052 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40093
VERSION E40093.1 GI:18627209
KEYWORDS JP 2000166577-A/7.
SOURCE unidentified
ORGANISM unidentified

```
            unclassified.
REFERENCE   1  (bases 1 to 2052)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 7 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/7
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2052).
FEATURES              Location/Qualifiers
     source           1. .2052
                      /organism="unidentified"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32644"
ORIGIN

  Query Match             99.8%;  Score 2048.8;  DB 6;  Length 2052;
  Best Local Similarity   99.9%;  Pred. No. 1e-309;
  Matches 2050;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780

Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900

Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960

Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020

Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080

Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140

Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200

Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260

Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320

Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380

Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440

Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500

Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560

Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620

Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680

Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
```

```
Db       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800

Qy       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860

Db       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860

Qy       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1920

Db       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1920

Qy       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980

Db       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980

Qy       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040

Db       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATCAC 2040

Qy       2041 ATTCTAAATATC 2052

Db       2041 ATTCTAAAGATC 2052


RESULT 3
E40091
LOCUS       E40091                  2056 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40091
VERSION     E40091.1  GI:18627207
KEYWORDS    JP 2000166577-A/5.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2056)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 5 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/5
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2056).
FEATURES             Location/Qualifiers
     source          1. .2056
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           99.3%;  Score 2038;  DB 6;  Length 2056;
  Best Local Similarity   99.8%;  Pred. No. 5e-308;
  Matches 2052;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Db        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Db        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Db        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Db        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Db        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780

Db        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780

Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840

Db        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900

Db        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900

Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960

Db        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960

Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020

Db        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020

Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080

Db       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080

Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140

Db       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140

Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200

Db       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200

Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260

Db       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
```

```
Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320

Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380

Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440

Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500

Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560

Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620

Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680

Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800

Qy       1801 GCATT----CTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTC 1856
             |||||    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCATTCTAGCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTC 1860

Qy       1857 AACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATT 1916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATT 1920

Qy       1917 TATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTA 1976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTA 1980

Qy       1977 TCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAA 2036
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAA 2040

Qy       2037 TAACATTCTAAATATC 2052
             ||||||||||||||||
Db       2041 TAACATTCTAAATATC 2056
```

```
RESULT 4
E40090
LOCUS       E40090                  2048 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40090
VERSION     E40090.1  GI:18627206
KEYWORDS    JP 2000166577-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2048)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 4 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/4
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2048).
FEATURES             Location/Qualifiers
     source          1. .2048
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN
```

Query Match 99.1%; Score 2034; DB 6; Length 2048;
Best Local Similarity 99.8%; Pred. No. 2.1e-307;
Matches 2048; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

```
Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||     ||
Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTC----GC 416

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 476

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 536

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 596

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 656

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 716

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 776
```

```
Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 836

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 896

Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 956

Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1016

Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1076

Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1136

Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1196

Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1197 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1256

Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1316

Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1317 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1376

Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1436

Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1437 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1496

Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1497 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1556

Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1557 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1616

Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1617 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1676

Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1677 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1736

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1737 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1796

Qy       1801 GCATTCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1797 GCATTCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1856
```

```
Qy       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1857 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATC 1916

Qy       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1917 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1976

Qy       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1977 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2036

Qy       2041 ATTCTAAATATC 2052
              ||||||||||||
Db       2037 ATTCTAAATATC 2048
```

```
RESULT 5
E40089
LOCUS       E40089                2048 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40089
VERSION     E40089.1  GI:18627205
KEYWORDS    JP 2000166577-A/3.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2048)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 3 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/3
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2048).
FEATURES             Location/Qualifiers
     source          1. .2048
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN
```

```
  Query Match             99.0%;  Score 2032.4;  DB 6;  Length 2048;
  Best Local Similarity   99.8%;  Pred. No. 3.7e-307;
  Matches 2047;  Conservative    0;  Mismatches    1;  Indels    4;  Gaps    1;

Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780

Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              |||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTG----CATTTCTTTAAAAGTAATG 896

Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 956

Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1016

Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1076

Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1136

Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1196

Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1197 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1256

Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1316

Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1317 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1376

Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1436

Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1437 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1496

Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1497 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1556

Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1557 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1616

Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1617 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1676

Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1677 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1736

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1737 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1796

Qy       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1797 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1856

Qy       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1857 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1916

Qy       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1917 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1976

Qy       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1977 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATCAC 2036

Qy       2041 ATTCTAAATATC 2052
              ||||||||||||
Db       2037 ATTCTAAATATC 2048
```

```
RESULT 6
AR076817
LOCUS       AR076817               2042 bp    DNA     linear   PAT 30-AUG-2000
DEFINITION  Sequence 2 from patent US 5959176.
ACCESSION   AR076817
VERSION     AR076817.1  GI:10003563
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2042)
  AUTHORS   Torikai,S. and Oeda,K.
  TITLE     Plant promoter and utilization thereof
  JOURNAL   Patent: US 5959176-A 2 28-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .2042
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           11.8%;  Score 241.2;  DB 6;  Length 2042;
  Best Local Similarity   72.8%;  Pred. No. 3.5e-28;
  Matches  367;  Conservative    0;  Mismatches  128;  Indels    9;  Gaps    4;
```

```
Qy       1554 AAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCG----TTCGGTCAAATG 1609
              |||| | |||||||||     ||     | | |     |||||  | |      | || | |
Db       1543 AAATTATTTATCTGAATGATAACATCTTTGTAAACAAAACTGGGCCAAATAGGACCATAA 1602

Qy       1610 GGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATT 1669
              |||||||  |||||||| | | | ||||||   ||  | | | ||    ||  ||      |
Db       1603 CCAAGTTCACGTGTATTCTAAAATGTTAATACTAACATGAGTATTTTCTT--TTCAAGGT 1660

Qy       1670 TTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGC 1729
                |  || |   | | | |   || |  |||  | |     |  |  |     |
Db       1661 ATAAGTTAATTCTTCAATCAATTAACTTTAAATTTGGACATTATTGAGCAACTTTATGCC 1720

Qy       1730 AGCACTAGACTGTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAAC 1789
                | |  | ||||| ||||| || |||||||||| || ||||||||||||||||||| |
Db       1721 CACGTTGTATTGTTTAAACAACGTTTGTCCGGTGTATATTTATGACCTTTCAACTCAAGC 1780

Qy       1790 TAGTGAAT-AATGCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACT 1848
              |||  | | ||||| |||||||||| ||||||| |||||||||||||||||||  |||||
Db       1781 TAGCCAGTGAATGCTTTCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACT 1840

Qy       1849 TTTCTTTCAACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATT 1908
              ||||||| ||| |||| ||||| ||| |||||||||| ||| |  ||||||| ||||| |
Db       1841 TTTCTTTTAACAGATTAGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACT 1898

Qy       1909 GTAATATTTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1899 ATAATATTTATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958

Qy       1969 CTCTACTATCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATT 2028
              |||||||||||||||||||||||| || |||||||||||||||||| |||||||| |||
Db       1959 CTCTACTATCATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATT 2018

Qy       2029 CTCAGCAATAACATTCTAAATATC 2052
                | ||||| ||||||||||  ||
Db       2019 TTTAGCAAAAACATTCTAAAGGTC 2042
```

RESULT 7
E15125

```
LOCUS       E15125                2042 bp    DNA     linear   PAT 28-JUL-1999
DEFINITION  Promoter.
ACCESSION   E15125
VERSION     E15125.1  GI:5709808
KEYWORDS    JP 1998052273-A/2.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2042)
  AUTHORS   Torikai,T. and Oita,K.
  TITLE     VEGETABLE PROMOTER AND ITS USE
  JOURNAL   Patent: JP 1998052273-A 2 24-FEB-1998;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L. (carrot)
            PN   JP 1998052273-A/2
            PD   24-FEB-1998
            PF   12-AUG-1996 JP 1996212680
            PI   TORIKAI TOSHIMI, OITA KENJI
            PC   C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
             strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2042
            FT                   /organism='Daucus carota L.'
            FT   promoter        1<. .<2042.
FEATURES                Location/Qualifiers
     source             1. .2042
                        /organism="unidentified"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32644"
ORIGIN
```

```
  Query Match            11.8%;  Score 241.2;  DB 6;  Length 2042;
  Best Local Similarity   72.8%;  Pred. No. 3.5e-28;
  Matches  367;  Conservative    0;  Mismatches  128;  Indels    9;  Gaps    4;

Qy       1554 AAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCG----TTCGGTCAAATG 1609
              |||| | |||||||||     ||     | | |     |||||  | |      | || | |
Db       1543 AAATTATTTATCTGAATGATAACATCTTTGTAAACAAAACTGGGCCAAATAGGACCATAA 1602

Qy       1610 GGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATT 1669
              |||||||  |||||||| | | | ||||||   ||  | | | ||    ||  ||      |
Db       1603 CCAAGTTCACGTGTATTCTAAAATGTTAATACTAACATGAGTATTTTCTT--TTCAAGGT 1660

Qy       1670 TTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGC 1729
                |  || |   | | | |   || |  |||  | |     |  |  |     |
Db       1661 ATAAGTTAATTCTTCAATCAATTAACTTTAAATTTGGACATTATTGAGCAACTTTATGCC 1720

Qy       1730 AGCACTAGACTGTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAAC 1789
                | |  | ||||| ||||| || |||||||||| || ||||||||||||||||||| |
Db       1721 CACGTTGTATTGTTTAAACAACGTTTGTCCGGTGTATATTTATGACCTTTCAACTCAAGC 1780

Qy       1790 TAGTGAAT-AATGCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACT 1848
              |||  | | ||||| |||||||||| ||||||| |||||||||||||||||||  |||||
Db       1781 TAGCCAGTGAATGCTTTCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACT 1840

Qy       1849 TTTCTTTCAACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATT 1908
              ||||||| ||| |||| ||||| ||| |||||||||| ||| |  ||||||| ||||| |
Db       1841 TTTCTTTTAACAGATTAGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACT 1898

Qy       1909 GTAATATTTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1899 ATAATATTTATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958

Qy       1969 CTCTACTATCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATT 2028
              |||||||||||||||||||||||| || |||||||||||||||||| |||||||| |||
Db       1959 CTCTACTATCATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATT 2018

Qy       2029 CTCAGCAATAACATTCTAAATATC 2052
                | ||||| ||||||||||  ||
Db       2019 TTTAGCAAAAACATTCTAAAGGTC 2042
```

RESULT 8
AR076816

```
LOCUS       AR076816                247 bp    DNA     linear   PAT 30-AUG-2000
DEFINITION  Sequence 1 from patent US 5959176.
ACCESSION   AR076816
VERSION     AR076816.1  GI:10003562
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 247)
  AUTHORS   Torikai,S. and Oeda,K.
  TITLE     Plant promoter and utilization thereof
  JOURNAL   Patent: US 5959176-A 1 28-SEP-1999;
FEATURES                Location/Qualifiers
     source             1. .247
                        /organism="unknown"
                        /mol_type="unassigned DNA"
ORIGIN
```

```
  Query Match             9.6%;  Score 196.6;  DB 6;  Length 247;
  Best Local Similarity   89.6%;  Pred. No. 4.7e-21;
  Matches  223;  Conservative    0;  Mismatches   24;  Indels    2;  Gaps    1;

Qy       1804 TTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGAT 1863
              |||||||||| ||||||| |||||||||||||||||  |||||||||| ||| |||
Db          1 TTCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACTTTTCTTTTAACAGAT 60

Qy       1864 TGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATCAAC 1923
              | ||||| ||| |||||||||| ||| |  ||||||| ||||| | ||||||||||||||
```

```
Db         61 TAGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACTATAATATTTATCAAC 118

Qy       1924 ACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 1983
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 178

Qy       1984 ATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAACATT 2043
              ||||||||| || |||||||||||||||||| |||||||| ||| | ||||| ||||||
Db        179 ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAAAACATT 238

Qy       2044 CTAAATATC 2052
              |||||  ||
Db        239 CTAAAGGTC 247


RESULT 9
E15124
LOCUS       E15124                  247 bp    DNA     linear   PAT 28-JUL-1999
DEFINITION  Promoter.
ACCESSION   E15124
VERSION     E15124.1  GI:5709807
KEYWORDS    JP 1998052273-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 247)
  AUTHORS   Torikai,T. and Oita,K.
  TITLE     VEGETABLE PROMOTER AND ITS USE
  JOURNAL   Patent: JP 1998052273-A 1 24-FEB-1998;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L. (carrot)
            PN   JP 1998052273-A/1
            PD   24-FEB-1998
            PF   12-AUG-1996 JP 1996212680
            PI   TORIKAI TOSHIMI, OITA KENJI
            PC   C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
             strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .247
            FT                   /organism='Daucus carota L.'
            FT                   /clone='pCR16G1-Xb'
            FT   promoter        1<. .<247.
FEATURES            Location/Qualifiers
     source         1. .247
                    /organism="unidentified"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32644"
ORIGIN

  Query Match             9.6%;  Score 196.6;  DB 6;  Length 247;
  Best Local Similarity   89.6%;  Pred. No. 4.7e-21;
  Matches  223;  Conservative    0;  Mismatches   24;  Indels    2;  Gaps    1;

Qy       1804 TTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGAT 1863
              |||||||||| ||||||| ||||||||||||||||||  |||||||||||| ||| |||
Db          1 TTCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACTTTTCTTTTAACAGAT 60

Qy       1864 TGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATCAAC 1923
              | ||||| ||| |||||||||| ||| |  ||||||| ||||| | ||||||||||||||
Db         61 TAGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACTATAATATTTATCAAC 118

Qy       1924 ACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 1983
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 178

Qy       1984 ATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAACATT 2043
              ||||||||| || |||||||||||||||||| |||||||| ||| | ||||| ||||||
Db        179 ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAAAACATT 238

Qy       2044 CTAAATATC 2052
              |||||  ||
Db        239 CTAAAGGTC 247


RESULT 10
AR146852
LOCUS       AR146852                246 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 2 from patent US 6218598.
ACCESSION   AR146852
VERSION     AR146852.1  GI:15110041
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246)
  AUTHORS   Ishige,F., Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter
  JOURNAL   Patent: US 6218598-A 2 17-APR-2001;
FEATURES            Location/Qualifiers
     source         1. .246
                    /organism="unknown"
                    /mol_type="unassigned DNA"
ORIGIN

  Query Match             9.5%;  Score 195.6;  DB 6;  Length 246;
  Best Local Similarity   89.5%;  Pred. No. 6.7e-21;
  Matches  222;  Conservative    0;  Mismatches   24;  Indels    2;  Gaps    1;

Qy       1805 TCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGATT 1864
              ||||||||| ||||||| ||||||||||||||||||  |||||||||||| ||| ||||
Db          1 TCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACTTTTCTTTTAACAGATT 60

Qy       1865 GGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATCAACA 1924
              |||||| ||| |||||||||| ||| |  ||||||| ||||| | ||||||||||||||
Db         61 AGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACTATAATATTTATCAACA 118

Qy       1925 CCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACA 1984
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACA 178

Qy       1985 TCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAACATTC 2044
              |||||||| || |||||||||||||||||| |||||||| ||| | ||||| |||||||
Db        179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAAAACATTC 238

Qy       2045 TAAATATC 2052
              ||||  ||
Db        239 TAAAGGTC 246


RESULT 11
E55065
LOCUS       E55065                  246 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter.
ACCESSION   E55065
VERSION     E55065.1  GI:18625251
KEYWORDS    JP 2000083679-A/2.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 246)
  AUTHORS   Ishige,I., Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter
  JOURNAL   Patent: JP 2000083679-A 2 28-MAR-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000083679-A/2
            PD   28-MAR-2000
            PF   12-JUL-1999 JP 1999197240
            PR
            PI   IKUHARU ISHIGE,SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N15/09,C12R1:91), PC
              (C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:91), PC
```

```
           (C12N5/00,C12R1:91)
           CC
           FH   Key             Location/Qualifiers
           FT   source          1. .246
           FT                   /organism='Daucus carota L.'.
FEATURES             Location/Qualifiers
     source          1. .246
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             9.5%;  Score 195.6;  DB 6;  Length 246;
  Best Local Similarity   89.5%;  Pred. No. 6.7e-21;
  Matches  222;  Conservative    0;  Mismatches   24;  Indels    2;  Gaps    1;

Qy       1805 TCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGATT 1864
              ||||||||| ||||||| |||||||||||||||||||  |||||||||||| ||| ||||
Db             1 TCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACTTTTCTTTTAACAGATT 60

Qy       1865 GGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATCAACA 1924
              |||||  ||| |||||||||| ||| |  ||||||| ||||| | ||||||||||||||||
Db            61 AGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACTATAATATTTATCAACA 118

Qy       1925 CCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACA 1984
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db           119 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACA 178

Qy       1985 TCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAACATTC 2044
              |||||||| || |||||||||||||||||| |||||||| ||| | ||||| |||||||
Db           179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAAAACATTC 238

Qy       2045 TAAATATC 2052
              ||||  ||
Db           239 TAAAGGTC 246


RESULT 12
DCA18706/c
LOCUS       DCA18706               8005 bp    DNA     linear   PLN 10-FEB-1999
DEFINITION  Daucus carota Inv*Dc5 gene.
ACCESSION   Y18706
VERSION     Y18706.1  GI:4454114
KEYWORDS    beta-fructofuranosidase; Inv*Dc5 gene; invertase; isoform.
SOURCE      Daucus carota (carrot)
  ORGANISM  Daucus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Apiales; Apiaceae; Apioideae; Scandiceae;
            Daucinae; Daucus.
REFERENCE   1
  AUTHORS   Sturm,A.
  TITLE     Molecular characterisation and functional analysis of
            sucrose-cleaving enzymes in carrot (Daucus carota L.)
  JOURNAL   J. Exp. Bot. 47, 1187-1192 (1996)
REFERENCE   2  (bases 1 to 8005)
  AUTHORS   Sturm,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) A. Sturm, Friedrich Miescher-Institut,
            Maulbeerstr. 66, CH-4058 Basel, SWITZERLAND
COMMENT     Related sequence X67163.
FEATURES             Location/Qualifiers
     source          1. .8005
                     /organism="Daucus carota"
                     /mol_type="genomic DNA"
                     /cultivar="Nantaise"
                     /db_xref="taxon:4039"
     gene            1. .8005
                     /gene="Inv*Dc5"
     5'UTR           1. .2666
                     /gene="Inv*Dc5"
     CDS             join(2767. .3160,3724. .3732,4656. .5515,5605. .5766,
                     5869. .6104,6190. .6277,6384. .6587)
                     /gene="Inv*Dc5"
                     /EC_number="3.2.1.26"
                     /note="soluble acid invertase"
                     /codon_start=1
                     /product="beta-fructofuranosidase, isoform II"
                     /protein_id="CAA77266.1"
                     /db_xref="GI:4454115"
                     /db_xref="GOA:Q42722"
                     /db_xref="TrEMBL:Q42722"
                     /translation="MEHPITISHYTPLPDGEHSPSLTTTNTAEQSSRRRSLTFVLLFS
                     SILAACLVMGTMVLFPNSGNEAVEKSTVVPEETVEVAPRGVAEGVSMKSFRRPALNAE
                     PPANFPWNSNVLSWQRSSFHFQPNQNWMNDPNGPLFYKGWYHLFYQYNPDGAIWGNKI
                     VWGHAVSSDLIHWKHLPVAMVTDHWYDVNGVWTGSATILPDGQIVMLYTGSTNESVQV
                     QNLAYPADPSDPLLIEWVKYPGNPVLVPPPGIDFKDFRDPTTAWRTPEGKWRLIIGSK
                     LNKTGISLVYDTVDFKNFTLLDGVLHAVHGTGMWECVDFYPVSKFGENGLDTSFDGVG
                     VKHVMKASLDDDRNDYYAIGTYDPVSGKWVPDNPELDVGIGLRYDYGIYYASKTFYDS
                     NKKRRVLWSWIKETDSEISDVRKGWASVQGIPRTILFDPKTGSNLLQWPVEEVNKLRL
                     NKTVFENVEINTGAVLPLEIGSGSQLDITAEFEVDKESLERVQETNEVYDCKNNGGSS
                     GRGALGPFGLLILADKDLSEQTPVYFYIAKGSGGNLRTFFCADHSRSSKAVDVDKEIY
                     GSVVPVLRGEKLTMRILVDHSIVESFSQGGRTCITSRVYPTKAIYNNAKVFLFNNATE
                     ARIIASLNIWQMNTAQRQTHFADLVI"
     exon            2767. .3160
                     /gene="Inv*Dc5"
                     /number=1
     intron          3161. .3723
                     /gene="Inv*Dc5"
                     /number=1
     exon            3724. .3732
                     /gene="Inv*Dc5"
                     /number=2
     intron          3733. .4655
                     /gene="Inv*Dc5"
                     /number=2
     exon            4656. .5515
                     /gene="Inv*Dc5"
                     /number=3
     intron          5516. .5604
                     /gene="Inv*Dc5"
                     /number=3
     exon            5605. .5766
                     /gene="Inv*Dc5"
                     /number=4
     intron          5767. .5868
                     /gene="Inv*Dc5"
                     /number=4
     exon            5869. .6104
                     /gene="Inv*Dc5"
                     /number=5
     intron          6105. .6189
                     /gene="Inv*Dc5"
                     /number=5
     exon            6190. .6277
                     /gene="Inv*Dc5"
                     /number=6
     intron          6278. .6383
                     /gene="Inv*Dc5"
                     /number=6
     exon            6384. .6587
                     /gene="Inv*Dc5"
                     /number=7
     3'UTR           6588. .8005
                     /gene="Inv*Dc5"
ORIGIN

  Query Match             8.5%;  Score 174;  DB 8;  Length 8005;
  Best Local Similarity   81.7%;  Pred. No. 7.8e-18;
  Matches  201;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy         17 CACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGACTTCTTCTTCTTTTGA 76
              |||   |||  |||||| ||    |||| || |||||||||| | ||||| ||||| ||||
Db       1285 CACTATGGGGGTGTTTGTTTCCCGGAAGTAGGAGCTGCTTCCGTCTTCTGCTTCTCTTGA 1226

Qy         77 CCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAATACTAACTTCTCTCTC 136
```

```
                || |||||| |||| ||| ||||  |||||||||||| || |||| ||| ||||||||||
Db       1225 CCCGTTTGTGTAAATAAGCAGAAGCATTTTTAAAAAGTTGAGAATGCTAGCTTCTCTCTC 1166

Qy        137 ACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACTTCTCATTTCTACTCCA 196
              ||| |||| |||||||||||||||||||||||| || |||||||||| ||||||||| |
Db       1165 ACAGCTTCTGCTTCTTTTCCAAACACTTTATTAATTTATTTACTTCTCACTTCTACTCTA 1106

Qy        197 CTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAA 256
              ||| | | | |||| |||||||||||||||||||||||||||||||||| | || |
Db       1105 CTTTTCTAATTTAAGTAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCCCTATGACT 1046

Qy        257 GATCAT 262
              || | |
Db       1045 GACCCT 1040


RESULT 13
BD188672
LOCUS       BD188672               2831 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Promotor and use thereof.
ACCESSION   BD188672
VERSION     BD188672.1  GI:32998411
KEYWORDS    JP 2003000252-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2831)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Promotor and use thereof
  JOURNAL   Patent: JP 2003000252-A 1 07-JAN-2003;
            SUMITOMO CHEMICAL CO LTD
COMMENT     OS   Daucus carota L. (carrot)
            PN   JP 2003000252-A/1
            PD   07-JAN-2003
            PF   31-MAY-2001 JP 2001164069
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/
            PC   00,C12N5/00
            CC   Promotor and use thereof
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2831).
FEATURES              Location/Qualifiers
     source           1. .2831
                      /organism="unidentified"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32644"
ORIGIN

  Query Match             7.9%;  Score 162.4;  DB 6;  Length 2831;
  Best Local Similarity   82.5%;  Pred. No. 6.1e-16;
  Matches  198;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy         22 AGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGACTTCTTCTTCTTTTGACCTGT 81
              |||||||||||| || |  ||| |||||||||||||||||||| ||| |||| ||| ||
Db       1716 AGGGCCTGTTTGTTTTATGGAATCAGAAGCTGCTTCTGACTTCTGCTTTTTTT-ACCCGT 1774

Qy         82 TTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAATACTAACTTCTCTCTCACAAC 141
              |||| |||| ||| ||||  | |||||| |||||| |||| ||| |||||||||||| |
Db       1775 TTGTGTAAAAAAGCAGAAGCACTTTTAAGAAGCTGAGAATGCTAGTTTCTCTCTCACAGC 1834

Qy        142 TTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACTTCTCATTTCTACTCCACTTCT 201
              ||| |||||||||||||||||||||| |||||   ||||||||| |||| || |||||||
Db       1835 TTCTGCTTCTTTTCCAAACACTTTATCAACTTACCTACTTCTCACTTCTGCTTCACTTCT 1894

Qy        202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAAGATCA 261
              || |||||||||||| ||| |||||||||  |||||||||||||| | |   ||  | |
Db       1895 TTACTATAAGCAAGAAGTCAATTCTTTTAAATTAACCCAAACGGCCCCTAAGTAATTTGA 1954


RESULT 14
BD188674
LOCUS       BD188674               2865 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Promotor and use thereof.
ACCESSION   BD188674
VERSION     BD188674.1  GI:32998413
KEYWORDS    JP 2003000252-A/3.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2865)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Promotor and use thereof
  JOURNAL   Patent: JP 2003000252-A 3 07-JAN-2003;
            SUMITOMO CHEMICAL CO LTD
COMMENT     OS   Daucus carota L. (carrot)
            PN   JP 2003000252-A/3
            PD   07-JAN-2003
            PF   31-MAY-2001 JP 2001164069
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/
            PC   00,C12N5/00
            CC   Promotor and use thereof
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2865).
FEATURES              Location/Qualifiers
     source           1. .2865
                      /organism="unidentified"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32644"
ORIGIN

  Query Match             7.9%;  Score 162.4;  DB 6;  Length 2865;
  Best Local Similarity   82.5%;  Pred. No. 6.1e-16;
  Matches  198;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy         22 AGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGACTTCTTCTTCTTTTGACCTGT 81
              |||||||||||| || |  ||| |||||||||||||||||||| ||| |||| ||| ||
Db       1716 AGGGCCTGTTTGTTTTATGGAATCAGAAGCTGCTTCTGACTTCTGCTTTTTTT-ACCCGT 1774

Qy         82 TTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAATACTAACTTCTCTCTCACAAC 141
              |||| |||| ||| ||||  | |||||| |||||| |||| ||| |||||||||||| |
Db       1775 TTGTGTAAAAAAGCAGAAGCACTTTTAAGAAGCTGAGAATGCTAGTTTCTCTCTCACAGC 1834

Qy        142 TTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACTTCTCATTTCTACTCCACTTCT 201
              ||| |||||||||||||||||||||| |||||   ||||||||| |||| || |||||||
Db       1835 TTCTGCTTCTTTTCCAAACACTTTATCAACTTACCTACTTCTCACTTCTGCTTCACTTCT 1894

Qy        202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAAGATCA 261
              || |||||||||||| ||| |||||||||  |||||||||||||| | |   ||  | |
Db       1895 TTACTATAAGCAAGAAGTCAATTCTTTTAAATTAACCCAAACGGCCCCTAAGTAATTTGA 1954


RESULT 15
DARGCHS2
LOCUS       DARGCHS2               4886 bp    DNA     linear   PLN 14-APR-2000
DEFINITION  Carrot gCHS2 gene for chalcone synthase.
ACCESSION   D16255
VERSION     D16255.1  GI:441168
KEYWORDS    chalcone synthase.
SOURCE      Daucus carota (carrot)
  ORGANISM  Daucus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Apiales; Apiaceae; Apioideae; Scandiceae;
            Daucinae; Daucus.
REFERENCE   1  (bases 1 to 4886)
  AUTHORS   Ozeki,Y., Davies,E. and Takeda,J.
  TITLE     Structure and expression of chalcone synthase gene in carrot
            suspension cultured cells regulated by 2,4-D
  JOURNAL   Plant Cell Physiol. 34, 1029-1037 (1993)
REFERENCE   2  (bases 1 to 4886)
  AUTHORS   Ozeki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAY-1993) Yoshihiro Ozeki, College of Arts and
            Sciences, The University of Tokyo, Department of Biology; Komaba,
            Meguro-ku, Tokyo 153, Japan (Tel:03-3467-1171(ex.253),
```

```
            Fax:03-3485-2904)
COMMENT     Submitted (14-MAY-1993) to DDBJ by:
            Yoshihiro Ozeki
            Department of Biol. College of
            Arts & Science, Univ. of Tokyo
            3-8-1 Komaba, Meguro-ku
            Tokyo 153
            Japan
            Phone:  03-3467-1171 x253
            Fax:    03-3485-2904.
FEATURES             Location/Qualifiers
     source          1. .4886
                     /organism="Daucus carota"
                     /mol_type="genomic DNA"
                     /cultivar="Kurodagosun"
                     /db_xref="taxon:4039"
                     /dev_stage="cultured cells"
     misc_feature    1957. .1962
                     /note="TACPyAT motif"
     misc_feature    2060. .2065
                     /note="G-box"
     misc_feature    2286. .2291
                     /note="G-box"
     misc_feature    2303. .2308
                     /note="TACPyAT motif"
     TATA_signal     2327. .2331
     prim_transcript 2358. .4519
                     /note="CHS2 mRNA and intron"
     exon            2358. .2607
                     /number=1
     gene            join(2430. .2607,3395. .4386)
                     /gene="gCHS2"
     CDS             join(2430. .2607,3395. .4386)
                     /gene="gCHS2"
                     /EC_number="2.3.1.74"
                     /codon_start=1
                     /product="chalcone synthase"
                     /protein_id="BAA03784.1"
                     /db_xref="GI:565639"
                     /translation="MVTVNEFRKAQRAEGPATVLAIGTATPPNCVDQSAYADYYFRIT
                     NSEDKPELKEKFRRMCEKSMINTRYMHLTEDLLKQNPSFCEYMASSLDARQDIVVNEV
                     PKLGKEAALRAIKEWGQPKSKITHLIFCTTSGVDMPGADFRLTKLLGLRPSVKRFMMY
                     QQGCFAGGTVLRLAKDLAENNKNARVLVVCSEITVITFRGPNDTHLDSLVGQALFGDG
                     AGAVIVGSDPVIGIEKPLFEIVSAAQTILPDSDGAIDGHLREVGLTFHLLKDVPGLIS
                     KNIRKSLVEAFKPLGISDWNSIFWIAHPGGPAILDQVETELSLKPEKLKSTRQVLRDY
                     GNMSSACVLFILDEMRKASAKDGHRTTGEGLDWGVLFGFGPGLTVETVVLHSVPT"
     intron          2608. .3394
                     /gene="gCHS2"
     exon            3395. .4519
                     /number=2
     polyA_signal    4501. .4506
ORIGIN

  Query Match             7.9%;  Score 162.4;  DB 8;  Length 4886;
  Best Local Similarity   82.5%;  Pred. No. 5.5e-16;
  Matches  198;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy         11 CTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGACTTCTTCTTC 70
              |||     |  |||||||||||||| ||    |||||||||||||||||| | ||||| |||
Db       1355 CTAGTAGATGTAGGGCCTGTTTGGATGGTGGAAGCAGAAGCTGCTTCCGGCTTCTGCTTT 1414

Qy         71 TTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAATACTAACTTC 130
              | |||||| |||||  |||||||||||||   | ||||||  ||| || ||||||||  ||||
Db       1415 TCTTGACCCGTTTGGGTAAAGAAGTAGAAGCACTTTTAAGAAGTTGAGAATACTAGCTTC 1474

Qy        131 TCTCTCACAACTTCCG-CTTCTTTTCCAAACACTTTATTAACTTTTTTACTTCTCATTTC 189
              ||||||| | |||| |  || |||||||||||||||||| | || ||||||| ||| |||
Db       1475 TCTCTCAGAGCTTCTGTTTTTTTTTCCAAACACTTTATTCATTTATTTACTTTTCACTTC 1534

Qy        190 TACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTC 249
              |||||||||||||  | |||||||||| |||||||||||| ||||||||||||||| |
Db       1535 TACTCCACTTCTTTAATTTAAGCAAGAAGTCACTTCTTTTAAACTAACCCAAACGGCCCC 1594
```

```
Search completed: December  6, 2004, 22:47:27
Job time : 9444.04 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 15:16:44 ; Search time 1134.79 Seconds
(without alignments)
9492.325 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgccctacagcaca.....gcaataacattctaaatatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		2052	100.0	2052	3	AAA37959	Aaa37959 Carrot pr
2		2048.8	99.8	2052	3	AAA37961	Aaa37961 Carrot pr
3		2034.8	99.2	2056	3	AAA37964	Aaa37964 Plasmid #
4		2030.8	99.0	2048	3	AAA37962	Aaa37962 Plasmid #
5		2030.8	99.0	2048	3	AAA37963	Aaa37963 Plasmid #
6		239.6	11.7	2042	2	AAV15144	Aav15144 New promo
7		196.6	9.6	247	2	AAV15143	Aav15143 New promo
8		195.6	9.5	246	3	AAZ49611	Aaz49611 Carrot CR
9		194.4	9.5	196	12	ADP07499	Adp07499 Carrot DN
10		162.4	7.9	2831	10	ADC56759	Adc56759 Carrot DN
11		162.4	7.9	2865	10	ADC56761	Adc56761 Carrot DN
12	c	119	5.8	140	3	AAZ49616	Aaz49616 Oligonucl
13		119	5.8	140	3	AAZ49615	Aaz49615 Oligonucl
14		99.4	4.8	8056	8	ABZ10246	Abz10246 Haematopo
15		91.8	4.5	8056	8	ABZ10100	Abz10100 Haematopo
16	c	91.6	4.5	8056	8	ABZ10246	Abz10246 Haematopo
17	c	88.6	4.3	851	3	AAA37960	Aaa37960 Carrot te
18	c	88.6	4.3	8056	8	ABZ10100	Abz10100 Haematopo
19		83	4.0	158001	12	ADL17884	Adl17884 Human pho
20	c	75	3.7	11745	6	ABK28332	Abk28332 DNA trans
21	c	71.2	3.5	8222	8	ACF62816	Acf62816 Colon can
22	c	71.2	3.5	11222	10	ADB54318	Adb54318 Pretreate
23	c	71	3.5	5930	6	ABL32517	Abl32517 Human imm
24		69.8	3.4	6352	6	ABK31340	Abk31340 Signal tr
25		69.8	3.4	6352	6	ABL70563	Abl70563 Chemicall
26		69.8	3.4	6352	6	AAS61235	Aas61235 Human gen
27	c	69.6	3.4	8222	8	ACF62794	Acf62794 Colon can
28	c	69.6	3.4	11222	10	ADB54190	Adb54190 Pretreate
29	c	69.4	3.4	16258	6	ABK40038	Abk40038 Human che
30	c	69.4	3.4	16258	6	ABL70376	Abl70376 Chemicall
31		69.2	3.4	6045	6	ABK31541	Abk31541 Signal tr
32		69.2	3.4	6045	6	ABL70624	Abl70624 Chemicall
33	c	68.2	3.3	15548	6	ABL34155	Abl34155 Human imm
34		68	3.3	11422	6	ABK39936	Abk39936 Human che
35		68	3.3	11422	6	ABL32218	Abl32218 Human imm
36	c	67.8	3.3	15732	4	AAS45388	Aas45388 Chemicall
37	c	67.8	3.3	15732	6	ABK28233	Abk28233 DNA trans
38		67.6	3.3	5984	6	ABQ66994	Abq66994 Human ang
39		67.6	3.3	6050	6	ABL34011	Abl34011 Human imm
40		67.4	3.3	5748	6	ABL33142	Abl33142 Human imm
41		67.4	3.3	18154	6	ABL32255	Abl32255 Human imm
42		67.2	3.3	5807	6	ABL33155	Abl33155 Human imm
43		67.2	3.3	6816	12	ADQ24856	Adq24856 Human sof
44	c	67	3.3	115218	8	ACA64845	Aca64845 Human HNR
45		66.6	3.2	419	8	ABX46069	Abx46069 Bovine ES

ALIGNMENTS

RESULT 1
AAA37959
ID AAA37959 standard; DNA; 2052 BP.
XX
AC AAA37959;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #1.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 69-70; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene

XX
SQ Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;

Query Match 100.0%; Score 2052; DB 3; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780

Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900

Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960

Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020

Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080

Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140

Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200

Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260

Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320

Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380

Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440

Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500

Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560

Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620

Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680

Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800

Qy       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860

Qy       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATC 1920

Qy       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980

Qy       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040

Qy       2041 ATTCTAAATATC 2052
              ||||||||||||
Db       2041 ATTCTAAATATC 2052
```

RESULT 2
AAA37961
ID AAA37961 standard; DNA; 2052 BP.
XX
AC AAA37961;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #2.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 78-79; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 736 A; 318 C; 317 G; 681 T; 0 U; 0 Other;

Query Match 99.8%; Score 2048.8; DB 3; Length 2052;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2050; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780

Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900

Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960

Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020

Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080

Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140

Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200

Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260

Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320

Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380

Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
```

```
Db       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440

Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500

Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560

Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620

Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680

Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800

Qy       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860

Qy       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATC 1920

Qy       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980

Qy       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATCAC 2040

Qy       2041 ATTCTAAATATC 2052
             |||||||| |||
Db       2041 ATTCTAAAGATC 2052
```

```
RESULT 3
AAA37964
ID   AAA37964 standard; DNA; 2056 BP.
XX
AC   AAA37964;
XX
DT   18-AUG-2000  (first entry)
XX
DE   Plasmid #3 DNA sequence used in mutation of promoter sequence.
XX
KW   Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS   Daucus carota.
XX
PN   WO200020613-A1.
XX
PD   13-APR-2000.
XX
PF   28-SEP-1999;   99WO-JP005303.
XX
PR   02-OCT-1998;   98JP-00281124.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Nishikawa S,  Oeda K;
XX
DR   WPI; 2000-303791/26.
XX
PT   New Plant promoters and terminators from Daucus carota L., useful in
PT   plant breeding, for e.g. controlling fertilities of plants.
XX
PS   Example 8; Page 74-76; 81pp; English.
XX
CC   This sequence represents a plasmid sequence used in a method for
CC   introducing a mutation into a carrot promoter. The invention relates to
CC   plant promoters and terminators from Daucus carota L. which are capable
CC   of expressing a gene of interest in plants. The invention also includes a
CC   chimeric gene characterized in that it comprises the promoter and a
CC   desired gene linked to each other in the form capable of functioning. A
CC   method of producing a transformant comprises introducing the promoter,
CC   the chimeric gene or a vector comprising the promoter and a desired gene
CC   or terminator sequence into a host cell. The plant promoters and
CC   terminators are useful in plant breeding, for e.g. fertilities of plants
CC   may be controlled by expressing, in the host cells, a sense or antisense
CC   gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ   Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;

  Query Match             99.2%;  Score 2034.8;  DB 3;  Length 2056;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2050;  Conservative    0;  Mismatches    2;  Indels    4;  Gaps    1;

Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
```

```
Qy       1801 GCATT----CTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTC 1856
              |||||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCATTCTAGCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTC 1860
Qy       1857 AACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATT 1916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATT 1920
Qy       1917 TATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTA 1976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTA 1980
Qy       1977 TCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAA 2036
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAA 2040
Qy       2037 TAACATTCTAAATATC 2052
              | |||||||||| |||
Db       2041 TCACATTCTAAAGATC 2056
```

RESULT 4
AAA37962
ID AAA37962 standard; DNA; 2048 BP.
XX
AC AAA37962;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #1 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 71-73; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match 99.0%; Score 2030.8; DB 3; Length 2048;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

```
Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780

Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              ||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTG----CATTTCTTTAAAAGTAATG 896

Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 956

Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1016

Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1076

Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1136

Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1196

Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1197 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1256

Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1316

Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1317 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1376

Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1436

Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1437 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1496

Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1497 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1556

Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1557 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1616

Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1617 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1676

Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1677 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1736

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1737 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1796

Qy       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1797 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1856

Qy       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1857 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1916

Qy       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1917 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1976

Qy       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1977 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATCAC 2036

Qy       2041 ATTCTAAATATC 2052
              |||||||| |||
Db       2037 ATTCTAAAGATC 2048
```

RESULT 5
AAA37963
ID AAA37963 standard; DNA; 2048 BP.
XX

AC AAA37963;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #2 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 73-74; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match 99.0%; Score 2030.8; DB 3; Length 2048;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

```
Qy          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGTGTGCCCTACAGCACATAGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60

Qy         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAA 120

Qy        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACT 180

Qy        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

Qy        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCCTCAATAAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300

Qy        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACAGGGTTATTTTTTAACGTGTCAACAAATTCTAATAATTTTACCTGGCCGGTGAACA 360

Qy        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTCGCATGC 420
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db        361 CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAACCAAATTC----GC 416

Qy        421 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 AGGACGACTTAGGTGAATACACATTGTACTGTGAGTCTTTAAACAAAGAACAAGTGGTTC 476

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 536

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 596

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 656

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 716

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 776

Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 836

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 896

Qy        901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 956

Qy        961 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 TTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC 1016

Qy       1021 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 TGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAA 1076

Qy       1081 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 AAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT 1136

Qy       1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1196

Qy       1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1197 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAG 1256

Qy       1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTG 1316

Qy       1321 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1317 AGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT 1376

Qy       1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCA 1436

Qy       1441 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1437 TAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATT 1496

Qy       1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1497 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAG 1556

Qy       1561 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1557 TTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATG 1616

Qy       1621 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1617 TGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGA 1676

Qy       1681 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1677 AATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACT 1736

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1737 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1796

Qy       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1797 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1856

Qy       1861 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1857 GATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATC 1916

Qy       1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1917 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1976

Qy       1981 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAAC 2040
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1977 CACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATCAC 2036

Qy       2041 ATTCTAAATATC 2052
              |||||||| |||
Db       2037 ATTCTAAAGATC 2048
```

RESULT 6
AAV15144
ID AAV15144 standard; DNA; 2042 BP.
XX
AC AAV15144;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX
PN EP824150-A2.
XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Torikai S, Oeda K;
XX
DR WPI; 1998-122310/12.
XX
PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
PS Claim 2; Page 15-16; 31pp; English.
XX
CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;

Query Match 11.7%; Score 239.6; DB 2; Length 2042;
Best Local Similarity 72.6%; Pred. No. 7.7e-35;
Matches 366; Conservative 0; Mismatches 129; Indels 9; Gaps 4;

```
Qy       1554 AAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCG----TTCGGTCAAATG 1609
              |||| | |||||||||     ||    | | |     ||||| |||      | || | |
Db       1543 AAATTATTTATCTGAATGATAACATCTTTGTAAACAAAACTGCGCCAAATAGGACCATAA 1602

Qy       1610 GGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATT 1669
              |||||| |||||||| | | | ||||||  ||  | | | ||   ||  ||      |
Db       1603 CCAAGTTCACGTGTATTCTAAAATGTTAATACTAACATGAGTATTTTCTT--TTCAAGGT 1660

Qy       1670 TTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGC 1729
              |  || |   | | | |  || |  ||| | |    | | |   | |
Db       1661 ATAAGTTAATTCTTCAATCAATTAACTTTAAATTTGGACATTATTGAGCAACTTTATGCC 1720

Qy       1730 AGCACTAGACTGTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAAC 1789
              | |  | ||||| ||||| || |||||||||| || ||||||||||||||||||| |
Db       1721 CACGTTGTATTGTTTAAACAACGTTTGTCCGGTGTATATTTATGACCTTTCAACTCAAGC 1780

Qy       1790 TAGTGAAT-AATGCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACT 1848
              |||  | | ||||| |||||||||| ||||||| |||||||||||||||||||  |||||
Db       1781 TAGCCAGTGAATGCTTTCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACT 1840

Qy       1849 TTTCTTTCAACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATT 1908
              ||||||| ||| |||| ||||| |||  ||||||||| ||| | |||||||  ||||| |
Db       1841 TTTCTTTTAACAGATTAGAATCGTTTCGTAAACTTTTAAAATT--AAAAAATACATTACT 1898

Qy       1909 GTAATATTTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1899 ATAATATTTATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958

Qy       1969 CTCTACTATCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATT 2028
              |||||||||||||||||||||| || |||||||||||||||||| ||||||| || |
Db       1959 CTCTACTATCATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACATAAT 2018

Qy       2029 CTCAGCAATAACATTCTAAATATC 2052
              | ||||| ||||||||||  ||
Db       2019 TTTAGCAAAAACATTCTAAAGGTC 2042
```

RESULT 7
AAV15143
ID AAV15143 standard; DNA; 247 BP.
XX
AC AAV15143;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX
PN EP824150-A2.

XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Torikai S, Oeda K;
XX
DR WPI; 1998-122310/12.
XX
PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
PS Claim 1; Page 14; 31pp; English.
XX
CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
SQ Sequence 247 BP; 88 A; 50 C; 21 G; 88 T; 0 U; 0 Other;

Query Match 9.6%; Score 196.6; DB 2; Length 247;
Best Local Similarity 89.6%; Pred. No. 5.1e-27;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

```
Qy       1804 TTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGAT 1863
             |||||||||| ||||||| ||||||||||||||||||  |||||||||||| ||| |||
Db          1 TTCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACTTTTCTTTTAACAGAT 60

Qy       1864 TGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAAATGCATTATTGTAATATTTATCAAC 1923
             | |||||  ||| ||||||||||  ||| |  |||||||  |||||  | ||||||||||||||
Db         61 TAGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACTATAATATTTATCAAC 118

Qy       1924 ACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 1983
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 178

Qy       1984 ATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAACATT 2043
             ||||||||| || |||||||||||||||||| |||||||| ||| | ||||| ||||||
Db        179 ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAAAACATT 238

Qy       2044 CTAAATATC 2052
             ||||| ||
Db        239 CTAAAGGTC 247
```

RESULT 8
AAZ49611
ID AAZ49611 standard; DNA; 246 BP.
XX
AC AAZ49611;
XX
DT 07-APR-2000 (first entry)
XX
DE Carrot CR16.3 fragment for synthesis of plant promoter.
XX
KW Synthetic DNA; plant promoter; CR16.3 fragment; carrot; transgenic plant;
KW soybean glycinin; stearoyl-ACP-desaturase gene;
KW male sterility-related gene; ds.
XX
OS Daucus carota.
XX
PN EP976832-A2.
XX
PD 02-FEB-2000.
XX
PF 13-JUL-1999; 99EP-00113732.
XX
PR 15-JUL-1998; 98JP-00200372.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Ishige F, Nishikawa S, Oeda K;
XX
DR WPI; 2000-128374/12.
XX
PT Novel promoter used to produce transgenic plants with higher expression
PT of a desired gene.
XX
PS Claim 1; Page 11-12; 24pp; English.
XX
CC The present sequence is a CR16.3 fragment from carrot genomic DNA. It is
CC used for synthesis of a plant promoter which comprises nucleotides 112-
CC 246, 54-246, or 1-246 of this sequence and a synthetic DNA. The promoter
CC is used for controlling the expression of a desired gene e.g. soybean
CC glycinin, stearoyl-ACP-desaturase and S-locus type specific RNase gene
CC (male sterility-related gene) in a host cell especially a microorganism
CC or a plant cell. The transformed plant cells can be used to produce
CC transgenic plants. The promoter is compact and therefore suitable for
CC higher expression of a desired gene in a particular tissue compared to
CC other host tissues
XX
SQ Sequence 246 BP; 88 A; 50 C; 21 G; 87 T; 0 U; 0 Other;

Query Match 9.5%; Score 195.6; DB 3; Length 246;
Best Local Similarity 89.5%; Pred. No. 7.8e-27;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

```
Qy       1805 TCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGATT 1864
             |||||||| ||||||| ||||||||||||||||||  |||||||||||| ||| ||||
Db          1 TCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACTTTTCTTTTAACAGATT 60

Qy       1865 GGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATCAACA 1924
              |||||  ||| ||||||||||  ||| |  |||||||  |||||  | ||||||||||||||
Db         61 AGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACTATAATATTTATCAACA 118

Qy       1925 CCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACA 1984
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACA 178

Qy       1985 TCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAACATTC 2044
             ||||||||| || |||||||||||||||||| |||||||| ||| | ||||| ||||||
Db        179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAAAACATTC 238

Qy       2045 TAAATATC 2052
             |||| ||
Db        239 TAAAGGTC 246
```

RESULT 9
ADP07499
ID ADP07499 standard; DNA; 196 BP.
XX
AC ADP07499;
XX
DT 29-JUL-2004 (first entry)
XX
DE Carrot DNA.
XX
KW Carrot; gene; ds; expression inducing promoter;
KW transcription start point; RNA polymerase II.
XX
OS Daucus carota.
XX
PN JP2004135597-A.
XX
PD 13-MAY-2004.
XX
PF 18-OCT-2002; 2002JP-00304115.

XX
PR 18-OCT-2002; 2002JP-00304115.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
DR WPI; 2004-360986/34.
XX
PT Expression inducing promoter, useful for expressing foreign gene,
PT comprising first DNA for detecting transcription start point and having
PT minimum promoter function, linked to second DNA having expression
PT inducing promoter function.
XX
PS Disclosure; SEQ ID NO 4; 61pp; Japanese.
XX
CC The invention relates to an expression inducing promoter comprising a
CC first DNA linked to a second DNA at its 5' terminus, where the first DNA
CC has a region which determines a transcription start point of RNA
CC polymerase II and has minimum promoter function, and the second DNA has
CC expression inducing promoter function in a plant cell. The expression
CC inducing promoter is useful for expressing a foreign gene. This sequence
CC represents carrot DNA used in the scope of the invention.
XX
SQ Sequence 196 BP; 66 A; 40 C; 19 G; 71 T; 0 U; 0 Other;

Query Match 9.5%; Score 194.4; DB 12; Length 196;
Best Local Similarity 99.5%; Pred. No. 1.3e-26;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1853 TTTCAACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAA 1912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTTCAACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAA 60

Qy       1913 TATTTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCT 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATTTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCT 120

Qy       1973 ACTATCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCA 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTATCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCA 180

Qy       2033 GCAATAACATTCTAAA 2048
             ||||| ||||||||||
Db        181 GCAATCACATTCTAAA 196
```

RESULT 10
ADC56759
ID ADC56759 standard; DNA; 2831 BP.
XX
AC ADC56759;
XX
DT 18-DEC-2003 (first entry)
XX
DE Carrot DNA that encodes an expression inducer type promoter.
XX
KW carrot; expression inducer type promoter; plant; vector; transformant;
KW ds.
XX
OS Daucus carota subsp. sativus.
XX
PN JP2003000252-A.
XX
PD 07-JAN-2003.
XX
PF 31-MAY-2001; 2001JP-00164069.
XX
PR 31-MAY-2001; 2001JP-00164069.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
DR WPI; 2003-472920/45.
XX
PT An expression inducer type promoter derived from DNA of Daucus carota L.
PT var. sativa DC. of 2831 bases and its analogues having the same
PT functions.
XX
PS Claim 1; SEQ ID NO 1; 16pp; Japanese.
XX
CC This invention relates to carrot DNA used as an expression inducer type
CC promoter. Specifically, it refers to promoter sequences derived from
CC plant DNA, preferably carrot, and functional mutants thereof that can be
CC used as expression inducers. Furthermore, the present invention describes
CC DNA, vectors, transformants and the process by which to prepare
CC transformants. This polynucleotide sequence is the DNA encoding the
CC carrot promoter of the invention.
XX
SQ Sequence 2831 BP; 950 A; 456 C; 395 G; 1030 T; 0 U; 0 Other;

Query Match 7.9%; Score 162.4; DB 10; Length 2831;
Best Local Similarity 82.5%; Pred. No. 1.4e-20;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

```
Qy         22 AGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGACTTCTTCTTCTTTTGACCTGT 81
              |||||||||||| || |  ||| ||||||||||||||||||||| ||| |||| ||| ||
Db       1716 AGGGCCTGTTTGTTTTATGGAATCAGAAGCTGCTTCTGACTTCTGCTTTTTTT-ACCCGT 1774

Qy         82 TTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAATACTAACTTCTCTCTCACAAC 141
              |||| |||| ||| ||||  | |||||| |||||| |||| |||  ||||||||||||  |
Db       1775 TTGTGTAAAAAAGCAGAAGCACTTTTAAGAAGCTGAGAATGCTAGTTTCTCTCTCACAGC 1834

Qy        142 TTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACTTCTCATTTCTACTCCACTTCT 201
              ||| |||||||||||||||||||||| |||||   |||||||||| |||| || |||||||
Db       1835 TTCTGCTTCTTTTCCAAACACTTTATCAACTTACCTACTTCTCACTTCTGCTTCACTTCT 1894

Qy        202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAAGATCA 261
              || |||||||||||||| ||| |||||||||  |||||||||||||| | |   ||  | |
Db       1895 TTACTATAAGCAAGAAGTCAATTCTTTTAAATTAACCCAAACGGCCCCTAAGTAATTTGA 1954
```

RESULT 11
ADC56761
ID ADC56761 standard; DNA; 2865 BP.
XX
AC ADC56761;
XX
DT 18-DEC-2003 (first entry)
XX
DE Carrot DNA encoding an expression inducer type promoter (SeqID 3).
XX
KW carrot; expression inducer type promoter; plant; vector; transformant;
KW ds.
XX
OS Daucus carota subsp. sativus.
XX
PN JP2003000252-A.
XX
PD 07-JAN-2003.
XX
PF 31-MAY-2001; 2001JP-00164069.
XX
PR 31-MAY-2001; 2001JP-00164069.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
DR WPI; 2003-472920/45.
XX
PT An expression inducer type promoter derived from DNA of Daucus carota L.
PT var. sativa DC. of 2831 bases and its analogues having the same
PT functions.
XX
PS Disclosure; SEQ ID NO 3; 16pp; Japanese.
XX
CC This invention relates to carrot DNA used as an expression inducer type
CC promoter. Specifically, it refers to promoter sequences derived from
CC plant DNA, preferably carrot, and functional mutants thereof that can be
CC used as expression inducers. Furthermore, the present invention describes

```
CC   DNA, vectors, transformants and the process by which to prepare
CC   transformants. This polynucleotide sequence is DNA encoding a carrot
CC   promoter (SeqID 3) of the invention.
XX
SQ   Sequence 2865 BP; 963 A; 461 C; 396 G; 1045 T; 0 U; 0 Other;

  Query Match             7.9%;  Score 162.4;  DB 10;  Length 2865;
  Best Local Similarity   82.5%;  Pred. No. 1.4e-20;
  Matches  198;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy         22 AGGGCCTGTTTGGTTGAGAGAAGCAGAAGCTGCTTCTGACTTCTTCTTCTTTTGACCTGT 81
              |||||||||||| || |  ||| ||||||||||||||||||||| ||| |||| ||| ||
Db       1716 AGGGCCTGTTTGTTTTATGGAATCAGAAGCTGCTTCTGACTTCTGCTTTTTTT-ACCCGT 1774

Qy         82 TTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTGCGAATACTAACTTCTCTCTCACAAC 141
              |||| |||| ||| ||||  | |||||| |||||| |||| |||  |||||||||||| |
Db       1775 TTGTGTAAAAAAGCAGAAGCACTTTTAAGAAGCTGAGAATGCTAGTTTCTCTCTCACAGC 1834

Qy        142 TTCCGCTTCTTTTCCAAACACTTTATTAACTTTTTTACTTCTCATTTCTACTCCACTTCT 201
              ||| |||||||||||||||||||| |||||   ||||||||| |||| || |||||||
Db       1835 TTCTGCTTCTTTTCCAAACACTTTATCAACTTACCTACTTCTCACTTCTGCTTCACTTCT 1894

Qy        202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAAGATCA 261
              || |||||||||||| ||| |||||||||  |||||||||||||| | |   ||  | |
Db       1895 TTACTATAAGCAAGAAGTCAATTCTTTTAAATTAACCCAAACGGCCCCTAAGTAATTTGA 1954


RESULT 12
AAZ49616/c
ID   AAZ49616 standard; DNA; 140 BP.
XX
AC   AAZ49616;
XX
DT   07-APR-2000  (first entry)
XX
DE   Oligonucleotide-4 for synthesis of CR16.1 fragment for plant promoter.
XX
KW   Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
KW   soybean glycinin; oligonucleotide; stearoyl-ACP-desaturase gene;
KW   male sterility-related gene; ss.
XX
OS   Daucus carota.
OS   Synthetic.
XX
PN   EP976832-A2.
XX
PD   02-FEB-2000.
XX
PF   13-JUL-1999;   99EP-00113732.
XX
PR   15-JUL-1998;   98JP-00200372.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Ishige F,  Nishikawa S,  Oeda K;
XX
DR   WPI; 2000-128374/12.
XX
PT   Novel promoter used to produce transgenic plants with higher expression
PT   of a desired gene.
XX
PS   Disclosure; Page 14; 24pp; English.
XX
CC   The present sequence is an oligonucleotide (- chain) used to prepare a
CC   CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
CC   fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
CC   construction of a plant promoter. The promoter is used for controlling
CC   the expression of a desired gene e.g. soybean glycinin, stearoyl-ACP-
CC   desaturase and S-locus type specific RNase gene (male sterility-related
CC   gene) in a host cell especially a microorganism or a plant cell. The
CC   transformed plant cells can be used to produce transgenic plants. The
CC   promoter is compact and therefore suitable for higher expression of a
CC   desired gene in a particular tissue compared to other host tissues
XX
SQ   Sequence 140 BP; 45 A; 17 C; 34 G; 44 T; 0 U; 0 Other;

  Query Match             5.8%;  Score 119;  DB 3;  Length 140;
  Best Local Similarity   92.6%;  Pred. No. 9.8e-13;
  Matches  125;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1918 ATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 1977
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        140 ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 81

Qy       1978 CATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAAT 2037
              |||||||||||||||| || |||||||||||||||||| |||||||| ||| | |||||
Db         80 CATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAA 21

Qy       2038 AACATTCTAAATATC 2052
              ||||||||||| ||
Db         20 AACATTCTAAAGGTC 6


RESULT 13
AAZ49615
ID   AAZ49615 standard; DNA; 140 BP.
XX
AC   AAZ49615;
XX
DT   07-APR-2000  (first entry)
XX
DE   Oligonucleotide-3 for synthesis of CR16.1 fragment for plant promoter.
XX
KW   Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
KW   soybean glycinin; oligonucleotide; stearoyl-ACP-desaturase gene;
KW   male sterility-related gene; ss.
XX
OS   Daucus carota.
OS   Synthetic.
XX
PN   EP976832-A2.
XX
PD   02-FEB-2000.
XX
PF   13-JUL-1999;   99EP-00113732.
XX
PR   15-JUL-1998;   98JP-00200372.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Ishige F,  Nishikawa S,  Oeda K;
XX
DR   WPI; 2000-128374/12.
XX
PT   Novel promoter used to produce transgenic plants with higher expression
PT   of a desired gene.
XX
PS   Disclosure; Page 13; 24pp; English.
XX
CC   The present sequence is an oligonucleotide (+ chain) used to prepare a
CC   CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
CC   fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
CC   construction of a plant promoter. The promoter is used for controlling
CC   the expression of a desired gene e.g. soybean glycinin, stearoyl-ACP-
CC   desaturase and S-locus type specific RNase gene (male sterility-related
CC   gene) in a host cell especially a microorganism or a plant cell. The
CC   transformed plant cells can be used to produce transgenic plants. The
CC   promoter is compact and therefore suitable for higher expression of a
CC   desired gene in a particular tissue compared to other host tissues
XX
SQ   Sequence 140 BP; 44 A; 34 C; 17 G; 45 T; 0 U; 0 Other;

  Query Match             5.8%;  Score 119;  DB 3;  Length 140;
  Best Local Similarity   92.6%;  Pred. No. 9.8e-13;
  Matches  125;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;
```

```
Qy       1918 ATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 1977
             ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db          5 ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 64

Qy       1978 CATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAAT 2037
             |||||||||||||| || |||||||||||||||||| |||||||| ||| | |||||
Db         65 CATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAA 124

Qy       2038 AACATTCTAAATATC 2052
             ||||||||||| ||
Db        125 AACATTCTAAAGGTC 139
```

RESULT 14
ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.
XX
AC ABZ10246;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwope I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 386; 117pp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 4.8%; Score 99.4; DB 8; Length 8056;
Best Local Similarity 44.5%; Pred. No. 6.3e-09;
Matches 624; Conservative 0; Mismatches 761; Indels 18; Gaps 5;

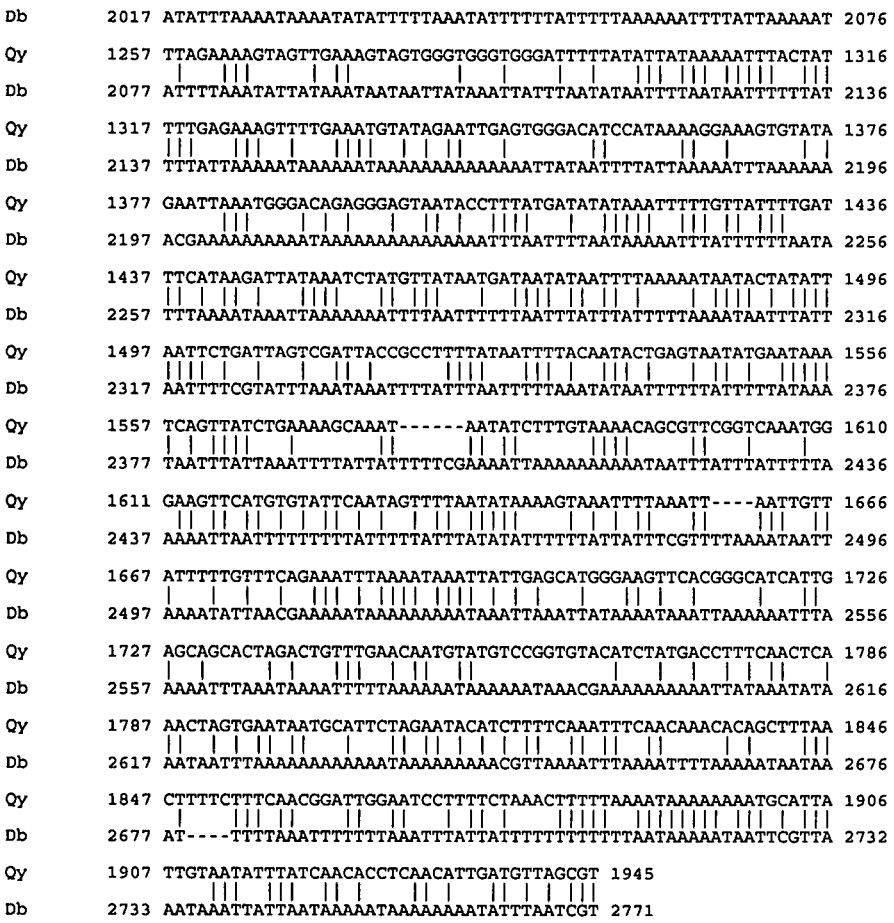
```
Qy        548 ATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGTAAAATTC 607
              ||   | | ||  | | ||| | || ||| | |||     |     |     || ||||||
Db       1975 ATTAATATTAATTTTTATTTAAAATTTAATAAAATTTTAAATATATTTAAAATAAAATAT 2034

Qy        608 CCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTT 667
              |||||  | ||| |   | || |   ||  | | | || |   ||    | |||  |
Db       2035 ATTTTTAAATATTTTTTATTTTTAAAAAATTTTATTAAAAATATTTTAAATATTATAAAT 2094

Qy        668 TTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCAC 727
              |  | ||   | |||    ||   |     |  | |     || |   |     |
Db       2095 AATAATTATAAATTATTTAATATAATTTTAATAATTTTTTATTTTATTAAAAATAAAAAA 2154

Qy        728 TGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAG 787
              | | ||| | | || |  || | | ||  |  |       | ||||| |  |||| || |
Db       2155 TAAAAAAAAAAAAAAATTATAATTTTATTAAAAATTTAAAAAAATGAAAAAAAAAATAAAA 2214

Qy        788 CTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAA 847
              |  ||  || ||| |  |||  || |    ||  | || |   |  |   |  || ||
Db       2215 AAAAAAAAAAATTTAATTTTAATAAAAATTTATTTTTTAATATTTAAAATAAATTAAAAA 2274

Qy        848 GAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTC-----TTTAAAAGTAATGT 901
              |  ||  |  | |   ||   |  || | |  | ||||        |||  |  |||
Db       2275 AATTTTAATTTTTTAATTTATTTATTTTTAAAATAATTTATTAATTTTTGTATTTAAATA 2334

Qy        902 AAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATAT 961
              ||  | || |||      | | | | ||  |     ||| || |  |  |  || || |||
Db       2335 AATTTTATTTAATTTTTAAATATAATTTTTTATTTTTATAAATAATTTATTAAATTTTAT 2394

Qy        962 TAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTACT 1021
              ||        | |||  | || |  |   |    | |  ||| ||      |   ||  |
Db       2395 TATTTTTTGAAAATTAAAAAAAAAAATAATTTATTTATTTTTAAAAATTAATT---TTTTT 2451

Qy       1022 GATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAA 1081
              |  ||| | |  || || |     |   | | ||||   |  |  |||||||
Db       2452 TTATTTTTATTTATATATTTTTTATTATTTTGTTTTAAAATAATTAAAATATTAATGAAA 2511

Qy       1082 AGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTA 1141
              |  | | ||  |   ||  | | ||  |  ||      |       | | ||    |
Db       2512 AATAAAAAAAAATAAATTAAATTATAAAATAAATTAAAAAATTTAAAAATTTAAATAAAA 2571

Qy       1142 TACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGA 1201
              |   ||   | |  | |    || ||||| ||||||   |  ||   ||||||  |
Db       2572 TTTTTAAAAAATAAAAAATAAATGAAAAAAAAAAATTATAAATATAAATAATTTAAAAAAA 2631

Qy       1202 AAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGA 1261
              || |  || ||||   | |||| |          |  ||||   ||||  || ||||
Db       2632 AAAAATAAAAAAAAAATGTTAAAATTTAAAATTTTAAAAATAATAAATTTTTAAATTTTTT 2691

Qy       1262 AAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGA 1321
              || |  | |  |  | |  |  | |  | | ||| | ||||    ||||   |
Db       2692 TAAATTTATTATTTTTTTTTTTTAATAAAAATAATTTGTTAAATAAAT----TATTAATA 2747

Qy       1322 GAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATT 1381
              |||        |||| | ||     |    |  || || ||  || | | | | | |
Db       2748 AAAATAAAAAAAATATTTAATTGTTTATTTTTAATAAATTTAAATGAATTTTGTTTATTA 2807

Qy       1382 AAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCAT 1441
              |||| |     |  |  |||   || |  | |||||| | ||  | |||||  | ||
Db       2808 AAATTGTTATTAATAATAAATGATATTTT--TTTATAAAATATTAAAAATTTGAAATTAT 2865

Qy       1442 AAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATTC 1501
              || | |||  | |  |  ||  | |  | || ||| ||||| | |  | | |  |
Db       2866 TTTATGAAAAAAATTTTGTTGAAAAAAATAAAAAATTTGAAAATGAAAAAAAAATGTTAA 2925

Qy       1502 TGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGT 1561
```

```
Db       2926 AATTTTATGAAATGTAAAATTTAAATGAATTATTATTGTTAAAAAAATAAAAAAAAATTA 2985

Qy       1562 TATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGT 1621

Db       2986 TTGAAAAATTGATATATTTTAAAATGATGATGTTTAAATTTTATAAATATAAATTAAAAA 3045

Qy       1622 GTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAA 1681

Db       3046 TAAAATAAATGTTATAATTGTGAATGAAATTTGAATTATTGTTATTTTTATGTGTAATTT 3105

Qy       1682 ATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACTG 1741

Db       3106 AATAAAAAATTAATTTTTTTAAAATGAATTTAATGATGATTGAAAAATTGTTTTATATAT 3165

Qy       1742 ---TTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATA 1798

Db       3166 AATTTTAAATAATAATATTTTTAATAAAAATTTTGTTGTTTGAATTGTAAATAAAATTTT 3225

Qy       1799 ATGCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAA 1858

Db       3226 AATTTATTATATTTTTTTTTAATAATTTGATTATTATTTTTAATTAATTGATTGTAAAAA 3285

Qy       1859 CGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTA 1918

Db       3286 AATTTTTGTTATTGTTGTTGTTATTTTTTTGAAAAAAAAAAAATATAATTGTAATGTAAA 3345

Qy       1919 TCAACACCTCAACATTGATGTTA 1941

Db       3346 AAAATGAATAAAATTTGAAATGA 3368
```

RESULT 15
ABZ10100
ID ABZ10100 standard; DNA; 8056 BP.
XX
AC ABZ10100;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #240.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwope I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 240; 117pp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;

Query Match 4.5%; Score 91.8; DB 8; Length 8056;
Best Local Similarity 44.2%; Pred. No. 1.6e-07;
Matches 627; Conservative 0; Mismatches 767; Indels 25; Gaps 5;

```
Qy        547 AATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGTAAAATT 606

Db       1358 ATTTTTTATTAAAAATTTATTTTTTTTTTATAAAAATAATATCGATTTTTTTTTTATTTT 1417

Qy        607 CCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACT 666

Db       1418 TTTATTTTTTTTTTTTAAAAAAAAAAATAAAAATTATTTTTTTAATAAAATTATATTAAA 1477

Qy        667 TTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCA 726

Db       1478 ATAATTAATTATTAAATTAATATTATATAAATATAAAATATAAATACGTTAAAAAAATTT 1537

Qy        727 CTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGA 786

Db       1538 TTAAATAAAATTATTTTTTTAAAAAATTATATTTAAATTTAATATTTTAATTTATATTTA 1597

Qy        787 GCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACA 846

Db       1598 TAATTTTAATATTTTAAAAAATCGAAATAAACGAATCGTAAAATTAAAAAAAATTATTTT 1657

Qy        847 AGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACT 906

Db       1658 AATTAATAC-GATAAAATTTTATTTTTATTAAAATTATAAAAATTAATTTAATATAATAA 1716

Qy        907 GATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATATTAAAC 966

Db       1717 TATATATTTATAATTTTAATTATTTAAAAAATTAAAATAAAAAAATTTTTTAAATTAAAA 1776

Qy        967 GAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCA 1026

Db       1777 AATTAAAAATTAATTACGTTATTATATTTTAATTTAATAATAAAATAAAATTTCGTTTTA 1836

Qy       1027 TCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAG 1086

Db       1837 AAAAAAAAAAAAAAAAAAAAAAAAAATTAATTTAATTAATTATTTTTATTTTTTTTATTAT 1896

Qy       1087 GACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACAT 1146

Db       1897 AAAAATAAAAATTTTATAAAAAAAAAAATAATTAAATTATAATATAAATAAAAAAAATAT 1956

Qy       1147 ATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAG 1206

Db       1957 AATTTTTTAAAATAAAATATTAATATTAATTTTTATTTAAAATTTAATAAAATTTTAAAT 2016

Qy       1207 A----------AAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGAT 1256
```



Search completed: December 6, 2004, 19:02:26
Job time : 1138.79 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:08:52 ; Search time 205.69 Seconds
(without alignments)
7090.965 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgccctacagcaca.....gcaataacattctaaatatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	241.2	11.8	2042	2	US-08-911-434A-2	Sequence 2, Appli
	2	196.6	9.6	247	2	US-08-911-434A-1	Sequence 1, Appli
	3	195.6	9.5	246	3	US-09-352-608-2	Sequence 2, Appli
	4	119	5.8	140	3	US-09-352-608-6	Sequence 6, Appli
c	5	119	5.8	140	3	US-09-352-608-7	Sequence 7, Appli
c	6	69.4	3.4	7218	1	US-08-232-463-14	Sequence 14, Appl
c	7	68.6	3.3	1141	4	US-09-806-708B-22	Sequence 22, Appl
	8	68.2	3.3	1141	4	US-09-806-708B-22	Sequence 22, Appl
	9	61.6	3.0	20674	4	US-09-641-638-651	Sequence 651, App
	10	61.6	3.0	20674	4	US-10-170-097-651	Sequence 651, App
c	11	60.2	2.9	19124	2	US-08-487-826B-13	Sequence 13, Appl
	12	59.6	2.9	7664	4	US-10-204-708-84	Sequence 84, Appl
c	13	59.6	2.9	53332	4	US-09-801-861-3	Sequence 3, Appli
c	14	59.6	2.9	53332	4	US-10-224-562-3	Sequence 3, Appli
c	15	58.6	2.9	5152	4	US-10-204-708-73	Sequence 73, Appl
c	16	58.2	2.8	5852	1	US-07-867-106-2	Sequence 2, Appli
c	17	58	2.8	8920	2	US-08-446-855A-1	Sequence 1, Appli
c	18	58	2.8	8920	3	US-09-150-741-1	Sequence 1, Appli
	19	58	2.8	10144	4	US-10-204-708-94	Sequence 94, Appl
c	20	57	2.8	640681	4	US-09-790-988-1	Sequence 1, Appli
	21	56.8	2.8	6583	4	US-10-204-708-25	Sequence 25, Appl
c	22	56.4	2.7	20674	4	US-09-641-638-651	Sequence 651, App
c	23	56.4	2.7	20674	4	US-10-170-097-651	Sequence 651, App
c	24	56	2.7	6113	4	US-10-204-708-13	Sequence 13, Appl
	25	55.8	2.7	658	3	US-08-998-416-595	Sequence 595, App
	26	55.8	2.7	6306	4	US-10-204-708-49	Sequence 49, Appl
	27	55.4	2.7	19124	2	US-08-487-826B-13	Sequence 13, Appl
	28	54.6	2.7	640681	4	US-09-790-988-1	Sequence 1, Appli
c	29	54.2	2.6	1055	4	US-09-806-708B-23	Sequence 23, Appl
	30	54.2	2.6	6317	4	US-10-204-708-11	Sequence 11, Appl
	31	53.8	2.6	1055	4	US-09-806-708B-23	Sequence 23, Appl
	32	53.4	2.6	6152	3	US-08-973-462-1	Sequence 1, Appli
	33	53.2	2.6	6156	4	US-10-204-708-59	Sequence 59, Appl
	34	53.2	2.6	168575	4	US-09-426-290-1	Sequence 1, Appli
c	35	53	2.6	636	3	US-08-998-416-1137	Sequence 1137, Ap
	36	53	2.6	53332	4	US-09-801-861-3	Sequence 3, Appli
	37	53	2.6	53332	4	US-10-224-562-3	Sequence 3, Appli
	38	52.8	2.6	782	3	US-08-998-416-224	Sequence 224, App
c	39	52.8	2.6	6243	2	US-09-056-075-1	Sequence 1, Appli
	40	52.8	2.6	319608	4	US-09-539-333D-1	Sequence 1, Appli
	41	52.8	2.6	319608	4	US-09-679-409-1	Sequence 1, Appli
	42	52.6	2.6	8920	2	US-08-446-855A-1	Sequence 1, Appli
	43	52.6	2.6	8920	3	US-09-150-741-1	Sequence 1, Appli
	44	52.4	2.6	636	3	US-08-998-416-1137	Sequence 1137, Ap
	45	52.4	2.6	1500	4	US-09-601-198-36	Sequence 36, Appl

ALIGNMENTS

```
RESULT 1
US-08-911-434A-2
; Sequence 2, Application US/08911434A
; Patent No. 5959176
;  GENERAL INFORMATION:
;    APPLICANT:  TORIKAI, Satomi
;    APPLICANT:  OEDA, Kenji
;    TITLE OF INVENTION:  PLANT PROMOTER AND UTILIZATION THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH, LLP
;      STREET:  P.O. BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VIRGINIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  22040
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/911,434A
;      FILING DATE:  12-AUG-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Raymond C.
;      REGISTRATION NUMBER:  21,066
;      REFERENCE/DOCKET NUMBER:  2185-0199P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)205-8000
;      TELEFAX:  (703)205-8050
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2042 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Daucus carota L.
;      INDIVIDUAL ISOLATE:  Kuroda Gosun
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  1..2042
US-08-911-434A-2

  Query Match            11.8%;  Score 241.2;  DB 2;  Length 2042;
  Best Local Similarity  72.8%;  Pred. No. 7.3e-42;
```

```
  Matches  367;  Conservative    0;  Mismatches  128;  Indels    9;  Gaps    4;

Qy       1554 AAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCG----TTCGGTCAAATG 1609
              |||| | |||||||||     ||     | | |     ||||| | |      | || | |
Db       1543 AAATTATTTATCTGAATGATAACATCTTTGTAAACAAAACTGGGCCAAATAGGACCATAA 1602

Qy       1610 GGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATT 1669
              ||||||| |||||||| | | | ||||||  ||  | | | ||    ||  ||       |
Db       1603 CCAAGTTCACGTGTATTCTAAAATGTTAATACTAACATGAGTATTTTCTT--TTCAAGGT 1660

Qy       1670 TTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGC 1729
              |   || |     |  |   |   || |   |||  | |      |  |   |     |
Db       1661 ATAAGTTAATTCTTCAATCAATTAACTTTAAATTTGGACATTATTGAGCAACTTTATGCC 1720

Qy       1730 AGCACTAGACTGTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAAC 1789
              |  |  | ||||| ||||| || |||||||||||| || ||||||||||||||||||| |
Db       1721 CACGTTGTATTGTTTAAACAACGTTTGTCCGGTGTATATTTATGACCTTTCAACTCAAGC 1780

Qy       1790 TAGTGAAT-AATGCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACT 1848
              |||  | | ||||| |||||||||| ||||||| ||||||||||||||||||  |||||
Db       1781 TAGCCAGTGAATGCTTTCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACT 1840

Qy       1849 TTTCTTTCAACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATT 1908
              ||||||| ||| |||| ||||| ||| |||||||||| ||| |  |||||||  ||||| |
Db       1841 TTTCTTTTAACAGATTAGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACT 1898

Qy       1909 GTAATATTTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1899 ATAATATTTATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958

Qy       1969 CTCTACTATCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATT 2028
              |||||||||||||||||||| || |||||||||||||||||| |||||||| |||
Db       1959 CTCTACTATCATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATT 2018

Qy       2029 CTCAGCAATAACATTCTAAATATC 2052
              | ||||| ||||||||||  ||
Db       2019 TTTAGCAAAAACATTCTAAAGGTC 2042


RESULT 2
US-08-911-434A-1
; Sequence 1, Application US/08911434A
; Patent No. 5959176
;  GENERAL INFORMATION:
;    APPLICANT:  TORIKAI, Satomi
;    APPLICANT:  OEDA, Kenji
;    TITLE OF INVENTION:  PLANT PROMOTER AND UTILIZATION THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH, LLP
;      STREET:  P.O. BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VIRGINIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  22040
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/911,434A
;      FILING DATE:  12-AUG-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Raymond C.
;      REGISTRATION NUMBER:  21,066
;      REFERENCE/DOCKET NUMBER:  2185-0199P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)205-8000
;      TELEFAX:  (703)205-8050
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Daucus carota L.
;      INDIVIDUAL ISOLATE:  Kuroda Gosun
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  1..247
US-08-911-434A-1

  Query Match             9.6%;  Score 196.6;  DB 2;  Length 247;
  Best Local Similarity   89.6%;  Pred. No. 1.1e-32;
  Matches  223;  Conservative    0;  Mismatches   24;  Indels    2;  Gaps    1;

Qy       1804 TTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGAT 1863
              |||||||||| ||||||| ||||||||||||||||||  |||||||||||| ||| |||
Db          1 TTCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACTTTTCTTTTAACAGAT 60

Qy       1864 TGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATCAAC 1923
              | ||||| ||| |||||||||| ||| |  |||||||  ||||| | ||||||||||||
Db         61 TAGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACTATAATATTTATCAAC 118

Qy       1924 ACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 1983
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 178

Qy       1984 ATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAACATT 2043
              ||||||||| || |||||||||||||||||| |||||||| ||| | ||||| ||||||
Db        179 ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAAAACATT 238

Qy       2044 CTAAATATC 2052
              |||||  ||
Db        239 CTAAAGGTC 247


RESULT 3
US-09-352-608-2
; Sequence 2, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
;  APPLICANT: ISHIGE, Fumiharu
;  APPLICANT: NISHIKAWA, Satomi
;  APPLICANT: OEDA, Kenji
;  TITLE OF INVENTION: Plant Promoter
;  FILE REFERENCE: 2185-0353P
;  CURRENT APPLICATION NUMBER: US/09/352,608
;  CURRENT FILING DATE: 1999-07-13
;  EARLIER APPLICATION NUMBER: 10-200372 JAPAN
;  EARLIER FILING DATE: 1998-07-15
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Daucus carota
US-09-352-608-2

  Query Match             9.5%;  Score 195.6;  DB 3;  Length 246;
  Best Local Similarity   89.5%;  Pred. No. 1.8e-32;
  Matches  222;  Conservative    0;  Mismatches   24;  Indels    2;  Gaps    1;

Qy       1805 TCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGATT 1864
              |||||||| ||||||| |||||||||||||||||  |||||||||||| ||| ||||
Db          1 TCTAGAATATATCTTTTGAAATTTCAACAAACACAGCACTAACTTTTCTTTTAACAGATT 60

Qy       1865 GGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTATCAACA 1924
              |||||  ||| |||||||||| ||| |  |||||||  ||||| | ||||||||||||||
Db         61 AGAATCGTTTCCTAAACTTTTAAAATT--AAAAAATACATTACTATAATATTTATCAACA 118
```

```
Qy       1925 CCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACA 1984
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACA 178

Qy       1985 TCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAATAACATTC 2044
              |||||||| || |||||||||||||||||| |||||||| ||| | ||||| |||||||
Db        179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAAAACATTC 238

Qy       2045 TAAATATC 2052
              ||||  ||
Db        239 TAAAGGTC 246


RESULT 4
US-09-352-608-6
; Sequence 6, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
;  APPLICANT: ISHIGE, Fumiharu
;  APPLICANT: NISHIKAWA, Satomi
;  APPLICANT: OEDA, Kenji
;  TITLE OF INVENTION: Plant Promoter
;  FILE REFERENCE: 2185-0353P
;  CURRENT APPLICATION NUMBER: US/09/352,608
;  CURRENT FILING DATE: 1999-07-13
;  EARLIER APPLICATION NUMBER: 10-200372 JAPAN
;  EARLIER FILING DATE: 1998-07-15
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 140
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-6

  Query Match             5.8%;  Score 119;  DB 3;  Length 140;
  Best Local Similarity   92.6%;  Pred. No. 2.1e-16;
  Matches  125;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1918 ATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 1977
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          5 ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 64

Qy       1978 CATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAAT 2037
              |||||||||||||||| || |||||||||||||||||| |||||||| ||| | |||||
Db         65 CATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAA 124

Qy       2038 AACATTCTAAATATC 2052
              ||||||||||| ||
Db        125 AACATTCTAAAGGTC 139


RESULT 5
US-09-352-608-7/c
; Sequence 7, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
;  APPLICANT: ISHIGE, Fumiharu
;  APPLICANT: NISHIKAWA, Satomi
;  APPLICANT: OEDA, Kenji
;  TITLE OF INVENTION: Plant Promoter
;  FILE REFERENCE: 2185-0353P
;  CURRENT APPLICATION NUMBER: US/09/352,608
;  CURRENT FILING DATE: 1999-07-13
;  EARLIER APPLICATION NUMBER: 10-200372 JAPAN
;  EARLIER FILING DATE: 1998-07-15
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 140
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-7

  Query Match             5.8%;  Score 119;  DB 3;  Length 140;
  Best Local Similarity   92.6%;  Pred. No. 2.1e-16;
  Matches  125;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1918 ATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 1977
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        140 ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 81

Qy       1978 CATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCAAT 2037
              |||||||||||||||| || |||||||||||||||||| |||||||| ||| | |||||
Db         80 CATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTTAGCAAA 21

Qy       2038 AACATTCTAAATATC 2052
              ||||||||||| ||
Db         20 AACATTCTAAAGGTC 6


RESULT 6
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             3.4%;  Score 69.4;  DB 1;  Length 7218;
  Best Local Similarity   4.6%;  Pred. No. 1.4e-05;
```

```
  Matches   16;  Conservative  212;  Mismatches  123;  Indels    0;  Gaps    0;

Qy       1051 ATTCAATGAGATAAAATATCTTACAATGAAAAGAAGGACAATGTCTCTTTGAAAAAACAA 1110
              | | || |||||| || |  ||   | :::::::::: :: :       ::::::: ::
Db       1461 AGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1402

Qy       1111 ATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATATGGATTGGACACGGAGACTAAGA 1170
              : ::: :      :       :::: : : : : : :::  ::: : :::::  ::::
Db       1401 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1342

Qy       1171 AAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGAAAGAAAAGTGGGTAAAGTAGCGG 1230
              :::: : : :::: :: : :::: :::::::::::::::::::: ::: :::: :: ::
Db       1341 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1282

Qy       1231 GACCCACCAATATATAATTGATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGG 1290
              ::   :  :: : : ::  :: :::   :::::::: ::  ::::: :: ::: ::: :::
Db       1281 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1222

Qy       1291 ATTTTTATATTATAAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTG 1350
              :   : :  : :::::   :  :    :::::::    :::: : : ::::  ::: :
Db       1221 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1162

Qy       1351 GGACATCCATAAAAGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAA 1401
              ::: :   : :::::::::: : : ::::  ::: :::: :::::::: ::
Db       1161 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1111


RESULT 7
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match             3.3%;  Score 68.6;  DB 4;  Length 1141;
  Best Local Similarity   13.1%;  Pred. No. 1.3e-05;
  Matches  141;  Conservative  392;  Mismatches  536;  Indels   10;  Gaps    5;

Qy        538 TTTTGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATA 597
              ||:|: : :|         :  ::   : |:: |||||::  :|::   :  ::     |
Db       1129 TTKTKTYKKANNNNNNNNGMGKDWNRMDATKWSATGTAWWTNHAKRGATMCWYWYWTGTN 1070

Qy        598 AGTAAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATT 657
              ::: :: : :   :|     ::  |:: ::|::::||:: || ::  :||::   ::  |
Db       1069 RRWCMRTYAMRTWYTRSNANWSCATKBMWWTMKWYATKYRTA-WYAMWCAWRNNNMWCAT 1011

Qy        658 CGTAATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAG 717
               |:|    |   :      :|   | ::|:     ::::| | | : | :||:| : ::
Db       1010 NGYAKSCATNNAMWYATTRWAAYAAAKWARWAGNNMRMYGAAAGNKWGCMAAMATMGBWW 951

Qy        718 ACTTGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATT 777
              | | |: |      :  :::: |:  : :      |::|| : : ||   :| : : :|:
Db        950 ADTAGKMCNNNNNNWTTDVRRMAM--KAKNNNNNNAYWTACYNRAATNNKMATHWMKWTH 893

Qy        778 AAAATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAG 837
               |: : : : :|:| ::|:        :   : ::| :::   ::             |

Db        892 GAHSKRRTRHHTRTCRRTKYNNNNNNARTVYWYHHAARRWMNAWWTRTNNNNNNNNNNAC 833

Qy        838 GTAAAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTA 897
              :  : :  :|:::: :                 :  :|    |:: ::       |::|:|
Db        832 RNTRTW--WABWKHSWCNNNNNNNNNNNNNNTWCHYTTANABBCYRANNNNAAARMARTC 775

Qy        898 ATGTAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTT 957
                 :||:|  : ::  : :: |: ::: ||::|   :  |    :|       : | |
Db        774 NNYMHAAVTTTHTDWCYKTWMNTWYWDMMTTMBTTTTRNMTTSTNMTNNNNNNMWACTNN 715

Qy        958 ATATTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACT 1017
                 : |: |:|   :  ::   : :     |  ::: : |  :|::  | ::   :::|
Db        714 NNNNMWKAYYAHATNNWGCWWNNTDARRTNNTTVMRRRWMTNTKTRWYSTTRR---HHYT 658

Qy       1018 TACTGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAAT 1077
               |                :| | | :: : : ::   :|:::   :::|   : :::|:
Db        657 GATNNNNNNNNNNNNNNNSCCTCTRMMTMRWTMKGDGMTVRKKVKWRDTTCTYVDVWAD 598

Qy       1078 GAAAAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAA 1137
              ::::: |   :|::  :    |:    |: :: :     ::  ::     ::   | :
Db        597 SWVWWYANWMRCRDVTYTRNNTYCKSYAHSYWYWSNNAMWYRRYSARNWSSMARWTTRNN 538

Qy       1138 TGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAA 1197
              :     :: :: |  : ::      :  : : : ::| :| | : ::    : : :|  ::
Db        537 WWMSGBVRMRWAGTMWWRHWNNNNTDTRYYWWWKRWARBTTTVYDSMCNAKSMWRGNNWR 478

Qy       1198 AAGAAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATT 1257
              |:::: | :|||::: |         :::   ::: :| |::          :| : :::
Db        477 AMKMWWAANNDAGAMDHWTYWMGNNTMWMRRAWKMMNMAWCRRAYCCNNNNNRACVWHKH 418

Qy       1258 TAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATT 1317
              :: :: ::::  ::||     ::  : ::      :|| |:: ::  :: :::   ::|
Db        417 KMWRWTWKYM--WKAACNNNNBKAMYMRVAWMMYSRDTTNTDWMMWTSDWBWHWYTVDYT 360

Qy       1318 TTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAG 1377
                :|        ::   |  : : :::       |   : :|: :|| :::|    ::
Db        359 MMRAWNNNNNNNWRBCKTTSWMWWMDHMNTHCTYGNNTWGSAYBMAAMSMWAAGASNBVT 300

Qy       1378 AATTAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATT 1437
                : :::|  |  :       :|::   :| |:   :: :: | ||::: : : :   ::
Db        299 YNWCWRMTYMGKTMTNNNNNNKAWYYRTKTVAWCNNRYYYDTAVWTBKRNYKYCYAYBWY 240

Qy       1438 TCATAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTA 1497
              ::::: :::: :::: :::     :::  : : :  :: :       :|:|:     : : :
Db        239 YBMYMGKHHWBWWRRABHRSWNMWWVKCRNKYMVSWHYHAMRYBKWABAVGCNNNWKDRM 180

Qy       1498 ATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAAT 1557
              |:::: ||    : :::|:: ::::   |:|  :  ::::: :::  : :: ::: |
Db        179 AHHHWCATNNNMMWWWYAYMHHMHKKGKAAWTNNKTABRDDHBAHVKTYWYWRYDYWCAM 120

Qy       1558 CAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTT 1616
              |:       : |:  ::  |::::    || :::::: | ::::   :: :: |:  |:
Db        119 CWMNAKAKVRTAMKHMWYYTDRYVSANNTGVRWMMMRWCMWWYSMNNRWYYRMGRKYTW 61


RESULT 8
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
```

```
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match             3.3%;  Score 68.2;  DB 4;  Length 1141;
  Best Local Similarity   12.6%;  Pred. No. 1.6e-05;
  Matches  137;  Conservative  405;  Mismatches  531;  Indels   13;  Gaps    5;

Qy        627 ATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAAGAGACTTA 686
Db         57 RKWTWARMYCKYRRWYNNKSRWWKGWYKKKWYBCANNTSBRYHARRWKDMKTAYBMTMTN 116

Qy        687 GGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGATAATTGTT 746
Db        117 KWGKTGWRHRYWRWRAMBDTVDHHYVTAMNNAWTTMCMMDKDDKRTRWWWKKNNNATGWD 176

Qy        747 AGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGTTACTAA 806
Db        177 DDTKYHMWNNNGCBTVTWMVRYKTDRDWSBKRMNYGMBWWKNWSYDVTYYWWVWDDMCKR 236

Qy        807 TAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAGAGCTTGCTGCTGTGTGTT 866
Db        237 KVRRWVRTRGRMRNYMV-AWBTAHRRRYNNGWTBAMAYRRWTMNNNNNNAKAMCKRAKYW 295

Qy        867 TAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGATCTAAAGCACATAGAAAT 926
Db        296 GWNRABVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDHKWWKWSAAMGVYWNNNNNNNW 355

Qy        927 TTAGTACAGGTTAAAACTTTTACAAGAATTTATATTAAACGAAAATCATTTTATAACATG 986
Db        356 TYKKARHBARWDWVWHSAWKKWHANAAHYSRKKWTBYKRKTMVNNNNGTTMWKRMWAWYW 415

Qy        987 TCTCTCGGCTGTC----ATTATAATAGGGATCACTTACTGATCATCCATTAAAACCTTGT 1042
Db        416 KMDMDWBGTYNNNNNGGRTYYGWTKNKKMWTYYKWKANNCKWRAWDHKTCTHNNTTWWKM 475

Qy       1043 TAAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGGACAATGTCTCTTTGA 1102
Db        476 KTYWNNCYWKSMTNGKSHRBAAAVYTWYMWWWRRYAHANNNNWDYWWKACTWYKYBVCSK 535

Qy       1103 AAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATATGGATTGGACACGGA 1162
Db        536 WWNNYAAWYTKSSWNYTSRYYRWKTNNSWRWRSDTRSMGRANNYARABHYGYKW-NTRWW 594

Qy       1163 GACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGAAAGAAAAGTGGGTAA 1222
Db        595 BWSHTWBHBRAGAAHYWMBMMYBAKCHCMKAWYKAKKYAGAGGSNNNNNNNNNNNNNNN 654

Qy       1223 AGTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTAGTTGAAAGTAGTGGG 1282
Db        655 ATCARDDYYAASRWYAMANAKWYYYKBAANNAYYTHANNWWGCWNNATDTRRTMWKNNNN 714

Qy       1283 TGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGA 1342
Db        715 NNAGTWKNNNNNNAKNASAAKNYAAAAVKAAKKHWRWANKWAMRGWHADAAABTTDKR-- 772

Qy       1343 ATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAAT 1402
Db        773 -NNGAYTKYTTTNNNNTYRGVVTNTAARDGWANNNNNNNNNNNNNNGWSDMWVTWWAYAN 831

Qy       1403 ACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATTATAAATCTATGTTAT 1462
Db        832 YGTNNNNNNNNNNAYAWWTNKWYYTTDDRWRBAYTNNNNNNRMAYYGAYADDYAYYMSDT 891

Qy       1463 AATGATAATATAATTTTAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGCCTT 1522
Db        892 CDAWMKWDATKMNNATTYNRGTAWRTNNNNNNMTMKTKYYBHAAWNNNNNNGKMCTAHTW 951

Qy       1523 TTATAATTTTACAATACTGAGTAATATGAATAAATCAGT----TATCTGAAAAGCAAATA 1578
Db        952 WVCKATKTTKGCWMNCTTTCRKYKNNCTWYTWMTTTRTTWYAATRWKTNNATGSMTRCNA 1011

Qy       1579 ATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAA 1638
Db       1012 TGWKNNNYWTGWKTRWTAYRMATRWMKAWWKVMATGSWNTNSYARWAYKTRAYKGWYYNA 1071

Qy       1639 TATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTAAAATAAATTATT 1698
Db       1072 CAWRWRWGKATCYMTDNAWWTACATSWMATHKYNWHMCKCNNNNNNNNNTMMRAMAMAAAN 1131

Qy       1699 GAGCAT 1704
Db       1132 CDGARY 1137


RESULT 9
US-09-641-638-651
; Sequence 651, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 651
;   LENGTH: 20674
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1123..3123
;   OTHER INFORMATION: 5'regulatory region
;   NAME/KEY: exon
;   LOCATION: 3124..3297
;   OTHER INFORMATION: exon 1
;   NAME/KEY: exon
;   LOCATION: 3871..4072
;   OTHER INFORMATION: exon 2
;   NAME/KEY: exon
;   LOCATION: 5552..5633
;   OTHER INFORMATION: exon 3
;   NAME/KEY: exon
;   LOCATION: 5758..5880
;   OTHER INFORMATION: exon 4
;   NAME/KEY: exon
;   LOCATION: 5996..6099
;   OTHER INFORMATION: exon 5
;   NAME/KEY: exon
;   LOCATION: 6349..6509
;   OTHER INFORMATION: exon 6
;   NAME/KEY: exon
;   LOCATION: 7379..7522
;   OTHER INFORMATION: exon 7
;   NAME/KEY: exon
;   LOCATION: 8645..8854
;   OTHER INFORMATION: exon 8
;   NAME/KEY: exon
;   LOCATION: 12254..12340
;   OTHER INFORMATION: exon 9
```

```
;  NAME/KEY: exon
;  LOCATION: 12854..13023
;  OTHER INFORMATION: exon 10
;  NAME/KEY: exon
;  LOCATION: 13308..13429
;  OTHER INFORMATION: exon 11
;  NAME/KEY: exon
;  LOCATION: 16567..16667
;  OTHER INFORMATION: exon 12
;  NAME/KEY: exon
;  LOCATION: 16775..16945
;  OTHER INFORMATION: exon 13
;  NAME/KEY: exon
;  LOCATION: 17063..17554
;  OTHER INFORMATION: exon 14
;  NAME/KEY: misc_feature
;  LOCATION: 17555..20674
;  OTHER INFORMATION: 3'regulatory region
;  NAME/KEY: allele
;  LOCATION: 1128
;  OTHER INFORMATION: 10-508-191  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 1182
;  OTHER INFORMATION: 10-508-245  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 1559
;  OTHER INFORMATION: 10-509-284  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 1570
;  OTHER INFORMATION: 10-509-295  : deletion of C
;  NAME/KEY: allele
;  LOCATION: 1827
;  OTHER INFORMATION: 10-510-173  : variable motif ATTTA or TTTTTT
;  NAME/KEY: allele
;  LOCATION: 2048
;  OTHER INFORMATION: 10-511-62  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 2323
;  OTHER INFORMATION: 10-511-337  : insertion of T
;  NAME/KEY: allele
;  LOCATION: 2341
;  OTHER INFORMATION: 10-512-36  : polymorphic base G or C
;  NAME/KEY: allele
;  LOCATION: 2623
;  OTHER INFORMATION: 10-512-318  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 2832
;  OTHER INFORMATION: 10-513-250  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 2844
;  OTHER INFORMATION: 10-513-262  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 2934
;  OTHER INFORMATION: 10-513-352  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 2947
;  OTHER INFORMATION: 10-513-365  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 3802
;  OTHER INFORMATION: 12-206-81  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 4062
;  OTHER INFORMATION: 10-343-231  : deletion of C
;  NAME/KEY: allele
;  LOCATION: 4088
;  OTHER INFORMATION: 12-206-366  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 4109
;  OTHER INFORMATION: 10-343-278  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 4170
;  OTHER INFORMATION: 10-343-339  : polymorphic base G or T
;  NAME/KEY: allele
;  LOCATION: 5903
;  OTHER INFORMATION: 10-346-23  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 6019
;  OTHER INFORMATION: 10-346-141  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 6141
;  OTHER INFORMATION: 10-346-263  : polymorphic base G or C
;  NAME/KEY: allele
;  LOCATION: 6183
;  OTHER INFORMATION: 10-346-305  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 6338
;  OTHER INFORMATION: 10-347-74  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 6375
;  OTHER INFORMATION: 10-347-111  : polymorphic base G or C
;  NAME/KEY: allele
;  LOCATION: 6429
;  OTHER INFORMATION: 10-347-165  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 6467
;  OTHER INFORMATION: 10-347-203  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 6484
;  OTHER INFORMATION: 10-347-220  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 6534
;  OTHER INFORMATION: 10-347-271  : polymorphic base A or T
;  NAME/KEY: allele
;  LOCATION: 6611
;  OTHER INFORMATION: 10-347-348  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 7668
;  OTHER INFORMATION: 10-348-391  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 8608
;  OTHER INFORMATION: 10-349-47  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 8658
;  OTHER INFORMATION: 10-349-97  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 8703
;  OTHER INFORMATION: 10-349-142  : polymorphic base G or C
;  NAME/KEY: allele
;  LOCATION: 8777
;  OTHER INFORMATION: 10-349-216  : deletion of CTG
;  NAME/KEY: allele
;  LOCATION: 8785
;  OTHER INFORMATION: 10-349-224  : polymorphic base G or T
;  NAME/KEY: allele
;  LOCATION: 8926
;  OTHER INFORMATION: 10-349-368  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 12171
;  OTHER INFORMATION: 10-350-72  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 12429
;  OTHER INFORMATION: 10-350-332  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 13341
;  OTHER INFORMATION: 10-507-170  : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 13492
;  OTHER INFORMATION: 10-507-321  : polymorphic base A or C
;  NAME/KEY: allele
;  LOCATION: 13524
;  OTHER INFORMATION: 10-507-353  : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 13535

Query Match             3.0%;  Score 61.6;  DB 4;  Length 20674;
Best Local Similarity   48.3%;  Pred. No. 0.00075;
```

```
  Matches  202;  Conservative    0;  Mismatches  214;  Indels    2;  Gaps    1;

Qy       1297 ATATTATAAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACAT 1356
              ||||| ||| | |||    ||   | |   |  | ||| |  | | |  | |   | |
Db      11084 ATATTTTAATATTTTTAATTTAATAAATTTTAATTTAATATTAAAATTAAAATTTTAATT 11143

Qy       1357 CCATAAAAGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAATACCTTTATGATATA 1416
                |||  |    |   | |  | || ||    | |     | |   |      | | ||
Db      11144 TAATATTAAATTAAAATTTTAATTTTAAATAAATAATTAAAATTTAAATAAATTTAAATT 11203

Qy       1417 TAAATTTTTGTTATTTTGATTTCATAAGATTATAAATCTATGTTATAATGATAATATAAT 1476
              |||| |||   | |  |  |     || |||  || | |   || ||||  |||  ||||
Db      11204 TAAAATTTAAATTTAATTTTAAATAAACATTTAAATTTTTAATTTTAATATTAAATTAAT 11263

Qy       1477 TTTAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAA 1536
              ||  ||  ||||| ||  |||| |    ||| |  | ||     ||   | | ||||
Db      11264 TT--AATTTAATATTAATTTAACTTAATTTAATTTAATATTAATTTAACTTAATTTAATT 11321

Qy       1537 TACTGAGTAATATGAATAAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGC 1596
              | | | || ||     || |    ||| |  ||    |||| ||  ||  |  ||
Db      11322 AATTTAATATTAAATTAAATTTTAATATTTTTAATTTTAATATTAATTTAATTTAATTTA 11381

Qy       1597 GTTCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAA 1656
              ||   |   |     |  |   | |   || ||| | || |||||  ||| ||||||||||
Db      11382 ATTTAATATTAATTAATTTAATTTTAATTTTAATTGCTTCAATATTAAATTAAATTTTAA 11441

Qy       1657 ATTAATTGTTATTTTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTTCA 1714
                |  || || |  |  |     ||| ||  |||  |  ||| |  ||   |  || |
Db      11442 TATTTTTATTTTAATATTAAATTAAAATTTTAATTTAAAATTAAATATTTAATATTAA 11499


RESULT 10
US-10-170-097-651
; Sequence 651, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE: GEN-T114XC2D1
;  CURRENT APPLICATION NUMBER: US/10/170,097
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: US 09/641,638
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: US 09/502,330
;  PRIOR FILING DATE: 2000-02-11
;  PRIOR APPLICATION NUMBER: US 60/133,200
;  PRIOR FILING DATE: 1999-05-07
;  PRIOR APPLICATION NUMBER: US 09/275,267
;  PRIOR FILING DATE: 1999-03-23
;  PRIOR APPLICATION NUMBER: US 60/119,917
;  PRIOR FILING DATE: 1999-02-12
;  NUMBER OF SEQ ID NOS: 1304
;  SOFTWARE: Patent.pm
; SEQ ID NO 651
;   LENGTH: 20674
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1123..3123
;   OTHER INFORMATION: 5'regulatory region
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 3124..3297
;   OTHER INFORMATION: exon 1
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 3871..4072
;   OTHER INFORMATION: exon 2
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 5552..5633
;   OTHER INFORMATION: exon 3
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 5758..5880
;   OTHER INFORMATION: exon 4
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 5996..6099
;   OTHER INFORMATION: exon 5
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 6349..6509
;   OTHER INFORMATION: exon 6
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 7379..7522
;   OTHER INFORMATION: exon 7
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 8645..8854
;   OTHER INFORMATION: exon 8
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 12254..12340
;   OTHER INFORMATION: exon 9
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 12854..13023
;   OTHER INFORMATION: exon 10
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 13308..13429
;   OTHER INFORMATION: exon 11
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 16567..16667
;   OTHER INFORMATION: exon 12
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 16775..16945
;   OTHER INFORMATION: exon 13
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: 17063..17554
;   OTHER INFORMATION: exon 14
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 17555..20674
;   OTHER INFORMATION: 3'regulatory region
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1128
;   OTHER INFORMATION: 10-508-191  : polymorphic base C or T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1182
;   OTHER INFORMATION: 10-508-245  : polymorphic base C or T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1559
;   OTHER INFORMATION: 10-509-284  : polymorphic base C or T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1570
;   OTHER INFORMATION: 10-509-295  : deletion of C
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1827
;   OTHER INFORMATION: 10-510-173  : variable motif ATTTA or TTTTTT
;   FEATURE:
```

; NAME/KEY: allele
; LOCATION: 2048
; OTHER INFORMATION: 10-511-62 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2323
; OTHER INFORMATION: 10-511-337 : insertion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2341
; OTHER INFORMATION: 10-512-36 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2623
; OTHER INFORMATION: 10-512-318 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2832
; OTHER INFORMATION: 10-513-250 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2844
; OTHER INFORMATION: 10-513-262 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2934
; OTHER INFORMATION: 10-513-352 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2947
; OTHER INFORMATION: 10-513-365 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 3802
; OTHER INFORMATION: 12-206-81 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4062
; OTHER INFORMATION: 10-343-231 : deletion of C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4088
; OTHER INFORMATION: 12-206-366 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4109
; OTHER INFORMATION: 10-343-278 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4170
; OTHER INFORMATION: 10-343-339 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 5903
; OTHER INFORMATION: 10-346-23 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6019
; OTHER INFORMATION: 10-346-141 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6141
; OTHER INFORMATION: 10-346-263 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6183
; OTHER INFORMATION: 10-346-305 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6338
; OTHER INFORMATION: 10-347-74 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6375
; OTHER INFORMATION: 10-347-111 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6429
; OTHER INFORMATION: 10-347-165 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6467
; OTHER INFORMATION: 10-347-203 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6484
; OTHER INFORMATION: 10-347-220 : polymorphic base A or G

Query Match 3.0%; Score 61.6; DB 4; Length 20674;
Best Local Similarity 48.3%; Pred. No. 0.00075;
Matches 202; Conservative 0; Mismatches 214; Indels 2; Gaps 1;

```
Qy       1297 ATATTATAAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACAT 1356
              ||||| ||| | |||     ||   | |    |  | ||| |  || | |  | ||   | |
Db      11084 ATATTTTAATATTTTTAATTTAATAAATTTTAATTTAATATTAAAATTAAAATTTTAATT 11143

Qy       1357 CCATAAAAGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAATACCTTTATGATATA 1416
                |||  |    |  | |  | || ||    | |     | |  |     |  | ||
Db      11144 TAATATTAAATTAAAATTTTAATTTTAAATAAATAATTAAAATTTAAATAAATTTAAATT 11203

Qy       1417 TAAATTTTTGTTATTTTGATTTCATAAGATTATAAATCTATGTTATAATGATAATATAAT 1476
              |||| |||   | |  |  |     || |||  || | |   || ||||  |||  ||||
Db      11204 TAAAATTTAAATTTAATTTTAAATAAACATTTAAATTTTTAATTTTAATATTAAATTAAT 11263

Qy       1477 TTTAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAA 1536
              ||  ||  ||||| ||  |||| |    ||| |  | ||     ||   | | ||||
Db      11264 TT--AATTTAATATTAATTTAACTTAATTTAATTTAATATTAATTTAACTTAATTTAATT 11321

Qy       1537 TACTGAGTAATATGAATAAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGC 1596
                | | | || ||     || |    ||| |  ||    |||| ||  ||  |  ||
Db      11322 AATTTAATATTAAATTAAATTTTAATATTTTTAATTTTAATATTAATTTAATTTAATTTA 11381

Qy       1597 GTTCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAA 1656
              ||   |   |   |  |   | |   || ||| | || ||||| ||| ||||||||||
Db      11382 ATTTAATATTAATTAATTTAATTTTAATTTTAATTGCTTCAATATTAAATTAAATTTTAA 11441

Qy       1657 ATTAATTGTTATTTTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTTCA 1714
                |  || || |  |  |     ||| ||  |||  |  ||| |  ||   |  || |
Db      11442 TATTTTTATTTTAATATTAAATTAAAATTTTAATTTAAAATTAAATATTTAATATTAA 11499
```

RESULT 11
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellems, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,826B
;      FILING DATE:  10-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH121.001CP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19124 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-487-826B-13

  Query Match             2.9%;  Score 60.2;  DB 2;  Length 19124;
  Best Local Similarity   49.2%;  Pred. No. 0.0014;
  Matches  271;  Conservative    0;  Mismatches  273;  Indels    7;  Gaps    4;

Qy       1162 AGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGAAAGAAAAGTGGGTA 1221
             | | | |  ||||   ||| | |||| |  | ||  |   || | ||| ||||   | |
Db       1438 ATAATTATGAAAAAAAATATATTAATATTTATTATTATTTAAAAAAAAAAAAAAGAGAGA 1379

Qy       1222 AAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTAGTTGAAAGTAGTGG 1281
             |         |   | || |||| ||  ||| ||       |  ||   ||  |||  |
Db       1378 ACAATTAAATA---AAAAAGATATTATATATATATACTATTAGATACCAGATTGTATAGT 1322

Qy       1282 GTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAG 1341
             | |  |  || ||| ||||| |||| || | || | |  || ||    | | ||
Db       1321 TTTCTTTAGTCTTGTATGTTATACAAATATA-TTTTATTAATAATTTACATATTATAATA 1263

Qy       1342 AATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAA 1401
               ||||          | |  | |    |  | |||| |  || ||          |  |
Db       1262 TTTTGATAACATATACTAATATATATATATATATATATATATATATATTTTAAATAAT 1203

Qy       1402 TACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATTATAAATCTATGTTA 1461
               |  |     | |||    | ||  |    | ||| ||||| || | |||   ||
Db       1202 AGCTATAGAATTTTATCTTATATTAATTGCATTATTACATAACATAACAAAATAATTAAT 1143

Qy       1462 TAATGATAATATAATTTTAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGCCT 1521
               || |   ||||| ||||||  ||| || |||||  ||| |||   ||    | |
Db       1142 TACTATATATATATTTTTAATTATATTATTATATATATTTTGAAGCGTTATACAACATTA 1083

Qy       1522 TTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCTGAAAAGCA-AATAAT 1580
               | |||| | | | ||||| |   |  || ||    | || |  ||   | |||  ||
Db       1082 TATATACTATGATAATACAGGCGATAATACATTTTTATTTAATATATAATTAATAATTAT 1023

Qy       1581 ATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAATA 1640
               || | | |||   |    |   | ||  |  |  | |||   | | ||  ||| |   |
Db       1022 ATATATATAAGTAATATATTATTTAATGTGTTACATTATGACAACTTAAATGTTGTGTGA 963

Qy       1641 TAAAAGTAAATTTT--AAATTAATTGTTATTTTTGTTTCAGAAATTTAAAATAAATTATT 1698
               ||| ||||  | |  |||||  |  |  | ||    | ||||||   || ||| |||
Db        962 AAAATGTAATGTATGAAAATTGTTATTATATATTAAAATAAAAATTTTTTATTAATAATT 903

Qy       1699 GAGCATGGGAA 1709
               |  ||| | |
Db        902 TATAATGTGCA 892


RESULT 12
US-10-204-708-84
; Sequence 84, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/204,708
;  CURRENT FILING DATE:  2003-05-06
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 84
;   LENGTH: 7664
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-84

  Query Match             2.9%;  Score 59.6;  DB 4;  Length 7664;
  Best Local Similarity   50.2%;  Pred. No. 0.0016;
  Matches  204;  Conservative    0;  Mismatches  194;  Indels    8;  Gaps    2;

Qy       1304 AAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAA 1363
               || |  ||    || |    |  ||||| |   | | || |     || |  |    |||
Db        408 AATAGATTTGATTTGTTTTGAGTTTTTGGAGAATTTTGATTAATAAGGTAGTTTTGAAAA 467

Qy       1364 AGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTT 1423
               ||||  ||  ||| | || |||||    || |  ||  |    | |  | | |||||||
Db        468 AGGATTGTAAATATATTTGAATGGAGGTGATGTGGTTTTTTTATAAGTTTTTTTAAATTT 527

Qy       1424 TTGTTATTTTGATTTC-----ATAAGATTATAAATCTATGTTATAATGATAATATAATTT 1478
               || ||||| |  |||      |||||| | |  || | | |||||||| |  || | |||
Db        528 TTTTTATTATCGTTTTAGTATATAAGAGTTTTTATTTTTATTATAATGTTGTTAGAGTTT 587

Qy       1479 TAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATA 1538
               |||   |  | | |      || |  ||| |   |           |  ||||||||   |
Db        588 TAAGTTTTTTTTTTTGCGTTTTTTCGTTATTATTTAGTTTGAAATGTTAATTTTATGGTT 647

Qy       1539 CTGAGTAATATGAATAAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGT 1598
               |   | |  |   |  |  |   |  | |     |   |  | ||||| ||  |
Db        648 TTTTTTTAGGTTTTGATTTAGGGGTTTGGTATGTTGTAGTTATTGGGTTTAAAAGAGTAT 707

Qy       1599 TCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAAT 1658
               |   |   |  ||  ||  ||| | ||  |||||  |  | |  || | | |||||
Db        708 TGTATTTTA---GAGTTTAGTGTTTTTTAGATAGTAGTGTTTTTTAAATTATTTTTACGA 764

Qy       1659 TAATTGTTATTTTTGTTTCAGAAATTTAAAATAAATTATTGAGCAT 1704
               |  |||||||||||  || | |    ||   | | |||||  || ||
Db        765 TTTTTGTTATTTTTAATTAAAATTCGTATTTTTATTTATTTAGAAT 810


RESULT 13
US-09-801-861-3/c
; Sequence 3, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
;  APPLICANT:  YAN, Chunhua et al.
;  TITLE OF INVENTION:  ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE:  CL001098
;  CURRENT APPLICATION NUMBER: US/09/801,861
;  CURRENT FILING DATE:  2001-03-09
;  NUMBER OF SEQ ID NOS:  10
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 53332
;   TYPE: DNA
;   ORGANISM: Human
US-09-801-861-3

  Query Match             2.9%;  Score 59.6;  DB 4;  Length 53332;
  Best Local Similarity   47.8%;  Pred. No. 0.0024;
  Matches  173;  Conservative    0;  Mismatches  189;  Indels    0;  Gaps    0;

Qy       1328 TTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGG 1387
              | | |||   | | |||  |      | ||  ||||||  | |       | |  |||||
Db      31673 TCTCAAAAAAAAAAAAATACATATATATATATATAAAAAAATATATATAATATATAAATAT 31614

Qy       1388 GACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATT 1447
                |      ||||  | ||  ||||| | || |    ||| |    |  ||||   |
Db      31613 ATAAATATATAAAATATATATAATATATAAATATATAAAATATATAATATATATAAATAT 31554

Qy       1448 ATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATTCTGATTA 1507
              |||||| ||| | |  ||   ||||||  |  | | |    |||| ||  | |    |
Db      31553 ATAAATATATATAAATATATAAATATATATAAATATAAATATATATAATATATATATAAA 31494

Qy       1508 GTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCTG 1567
                ||     | |   ||| |   || |   | |||| |||| |  |  ||| |
Db      31493 TATATATAAAATATATATAAATATATAAAAATATAAAAATATAAATATAATAAATATATA 31434

Qy       1568 AAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATTC 1627
              ||||   |  ||||| |   || |      |      ||||   |  |       ||
Db      31433 AAAATATATAAATATATAAATATATACAAATATATATAAATATATAAATATATAAAATAT 31374

Qy       1628 AATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTAA 1687
              ||   | | | |||| || || || | ||||  ||   |||   | | |   ||| | |
Db      31373 AAATATATAACTATATAAATATATATAAAATATATAAATATAAATATATATAAAAATATA 31314

Qy       1688 AA 1689
                |
Db      31313 TA 31312


RESULT 14
US-10-224-562-3/c
; Sequence 3, Application US/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
;  APPLICANT: YAN, Chunhua et al.
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001098DIV
;  CURRENT APPLICATION NUMBER: US/10/224,562
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 53332
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-224-562-3

  Query Match             2.9%;  Score 59.6;  DB 4;  Length 53332;
  Best Local Similarity   47.8%;  Pred. No. 0.0024;
  Matches  173;  Conservative    0;  Mismatches  189;  Indels    0;  Gaps    0;

Qy       1328 TTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGG 1387
              | | |||   | | |||  |      | ||  ||||||  | |       | |  |||||
Db      31673 TCTCAAAAAAAAAAAAATACATATATATATATATAAAAAAATATATATAATATATAAATAT 31614

Qy       1388 GACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATT 1447
                |      ||||  | ||  ||||| | || |    ||| |    |  ||||   |
Db      31613 ATAAATATATAAAATATATATAATATATAAATATATAAAATATATAATATATATAAATAT 31554

Qy       1448 ATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATTCTGATTA 1507
              |||||| ||| | |  ||   ||||||  |  | | |    |||| ||  | |    |
Db      31553 ATAAATATATATAAATATATAAATATATATAAATATAAATATATATAATATATATATAAA 31494

Qy       1508 GTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCTG 1567
                ||     | |   ||| |   || |   | |||| |||| |  |  ||| |
Db      31493 TATATATAAAATATATATAAATATATAAAAATATAAAAATATAAATATAATAAATATATA 31434

Qy       1568 AAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATTC 1627
              ||||   |  ||||| |   || |      |      ||||   |  |       ||
Db      31433 AAAATATATAAATATATAAATATATACAAATATATATAAATATATAAATATATAAAATAT 31374

Qy       1628 AATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTAA 1687
              ||   | | | |||| || || || | ||||  ||   |||   | | |   ||| | |
Db      31373 AAATATATAACTATATAAATATATATAAAATATATAAATATAAATATATATAAAAATATA 31314

Qy       1688 AA 1689
                |
Db      31313 TA 31312


RESULT 15
US-10-204-708-73/c
; Sequence 73, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/204,708
;  CURRENT FILING DATE:  2003-05-06
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 73
;   LENGTH: 5152
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-73

  Query Match             2.9%;  Score 58.6;  DB 4;  Length 5152;
  Best Local Similarity   49.9%;  Pred. No. 0.0024;
  Matches  208;  Conservative    0;  Mismatches  199;  Indels   10;  Gaps    2;

Qy       1400 AATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATTATAAATCTATGT 1459
              ||  | |   |  ||||  | ||||   | || | |  |   |  |||||||  ||
Db       2063 AAACCTTCACTCCTATAATATTTTTATATCTTCTCAAATATAACTTTTATAAACCTTCAA 2004

Qy       1460 TATAATGATAATATAATTTTAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGC 1519
              | | | | | |  | || || ||  || |||| | |  ||  | |  | ||     |
Db       2003 AACATTAAAATTTAATTTAAAATAAATTTAATATACTTAAACTTTTCACCCAAT-----C 1949

Qy       1520 CTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCTGAAAAGCAAATAA 1579
              | || |||||||| || ||  | | |    ||  |||  |  ||  | |||| ||| | |
Db       1948 CCTTAATAATTTTATAAAACCAAATCACCAAAAATAATAATATAAATAAAAAACAACTTA 1889

Qy       1580 TATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAAT 1639
              || ||      |  | | |   |||||      |    ||    | |||| |   | | |
Db       1888 AATTTT-----ATAATCTTAAAATCAAAAAACTAAAAAATTAAAAATCAAAATAATCATT 1834

Qy       1640 ATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTAAAATAAATTATTG 1699
```

```
Db       1833 AAAAAAAATATTTTTAATTCTATAAATACTACATTTTCATAAAATTAACTAAAAAAATTA 1774

Qy       1700 AGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACTGTTTGAACAATGTATGTCC 1759

Db       1773 ATTATAAAAAATTAACATAAAAAAATTAATTATAATATACTTTATCCGTAAAACATCTAT 1714

Qy       1760 GGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAATGCATTCTAGAATACAT 1816

Db       1713 TTTTAATCTTTATAATTTATCAATAAAACTATAAACAAAAAAAATTCCTAACCAAAT 1657
```

Search completed: December 6, 2004, 22:52:26
Job time : 208.69 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 19:02:39 ; Search time 1212.45 Seconds
(without alignments)
9300.077 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgccctacagcaca.....gcaataacattctaaatatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3695051 seqs, 2747533894 residues

Total number of hits satisfying chosen parameters: 7390102

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		99.4	4.8	8056	18	US-10-473-126-386	Sequence 386, App
c 2		93	4.5	3673778	15	US-10-312-841-1	Sequence 1, Appli
3		91.8	4.5	8056	18	US-10-473-126-240	Sequence 240, App
c 4		91.6	4.5	8056	18	US-10-473-126-386	Sequence 386, App
c 5		88.6	4.3	8056	18	US-10-473-126-240	Sequence 240, App
6		83	4.0	158001	16	US-10-211-179-11	GENERAL INFORMATI
c 7		75	3.7	11745	15	US-10-240-453-206	Sequence 206, App
c 8		71	3.5	5930	15	US-10-311-455-490	Sequence 490, App
9		70.6	3.4	3673778	15	US-10-312-841-1	Sequence 1, Appli
c 10		70	3.4	3252	13	US-10-027-632-113786	Sequence 113786,
c 11		70	3.4	3252	13	US-10-027-632-113787	Sequence 113787,
c 12		70	3.4	3252	13	US-10-027-632-113788	Sequence 113788,
c 13		70	3.4	3252	15	US-10-027-632-113786	Sequence 113786,
c 14		70	3.4	3252	15	US-10-027-632-113787	Sequence 113787,
c 15		70	3.4	3252	15	US-10-027-632-113788	Sequence 113788,
16		69.8	3.4	6352	16	US-10-221-613-195	Sequence 195, App
c 17		69.4	3.4	16258	16	US-10-257-166-120	Sequence 120, App
18		69	3.4	3673778	15	US-10-312-841-2	Sequence 2, Appli
c 19		68.6	3.3	49979	17	US-10-741-601-5746	Sequence 5746, Ap
c 20		68.2	3.3	15548	15	US-10-311-455-2128	Sequence 2128, Ap
21		68	3.3	11422	15	US-10-311-455-191	Sequence 191, App
22		68	3.3	11422	16	US-10-257-166-17	Sequence 17, Appl
c 23		67.8	3.3	15732	14	US-10-239-676-95	Sequence 95, Appl
c 24		67.8	3.3	15732	15	US-10-240-453-107	Sequence 107, App
25		67.6	3.3	5984	17	US-10-433-793-24	Sequence 24, Appl
26		67.6	3.3	6050	15	US-10-311-455-1984	Sequence 1984, Ap
27		67.4	3.3	5748	15	US-10-311-455-1115	Sequence 1115, Ap
28		67.4	3.3	18154	15	US-10-311-455-228	Sequence 228, App
29		67.2	3.3	5807	15	US-10-311-455-1128	Sequence 1128, Ap
30		67	3.3	2053	13	US-10-027-632-97533	Sequence 97533, A
31		67	3.3	2053	15	US-10-027-632-97533	Sequence 97533, A
32		66.6	3.2	419	9	US-09-960-352-11234	Sequence 11234, A
33		66.6	3.2	6079	15	US-10-311-455-394	Sequence 394, App
c 34		66.6	3.2	6292	16	US-10-221-714A-461	Sequence 461, App
c 35		66.4	3.2	18624	15	US-10-311-455-1676	Sequence 1676, Ap
36		66.2	3.2	1501	18	US-10-473-126-328	Sequence 328, App
37		66.2	3.2	3252	13	US-10-027-632-113786	Sequence 113786,
38		66.2	3.2	3252	13	US-10-027-632-113787	Sequence 113787,
39		66.2	3.2	3252	13	US-10-027-632-113788	Sequence 113788,
40		66.2	3.2	3252	15	US-10-027-632-113786	Sequence 113786,
41		66.2	3.2	3252	15	US-10-027-632-113787	Sequence 113787,
42		66.2	3.2	3252	15	US-10-027-632-113788	Sequence 113788,
c 43		66.2	3.2	6175	15	US-10-311-455-1280	Sequence 1280, Ap
44		65.8	3.2	15161	16	US-10-221-613-386	Sequence 386, App
c 45		65.2	3.2	34722	17	US-10-322-281-700	Sequence 700, App

ALIGNMENTS

RESULT 1
US-10-473-126-386
; Sequence 386, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 386
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-386

Query Match 4.8%; Score 99.4; DB 18; Length 8056;
Best Local Similarity 44.5%; Pred. No. 9.7e-08;
Matches 624; Conservative 0; Mismatches 761; Indels 18; Gaps 5;

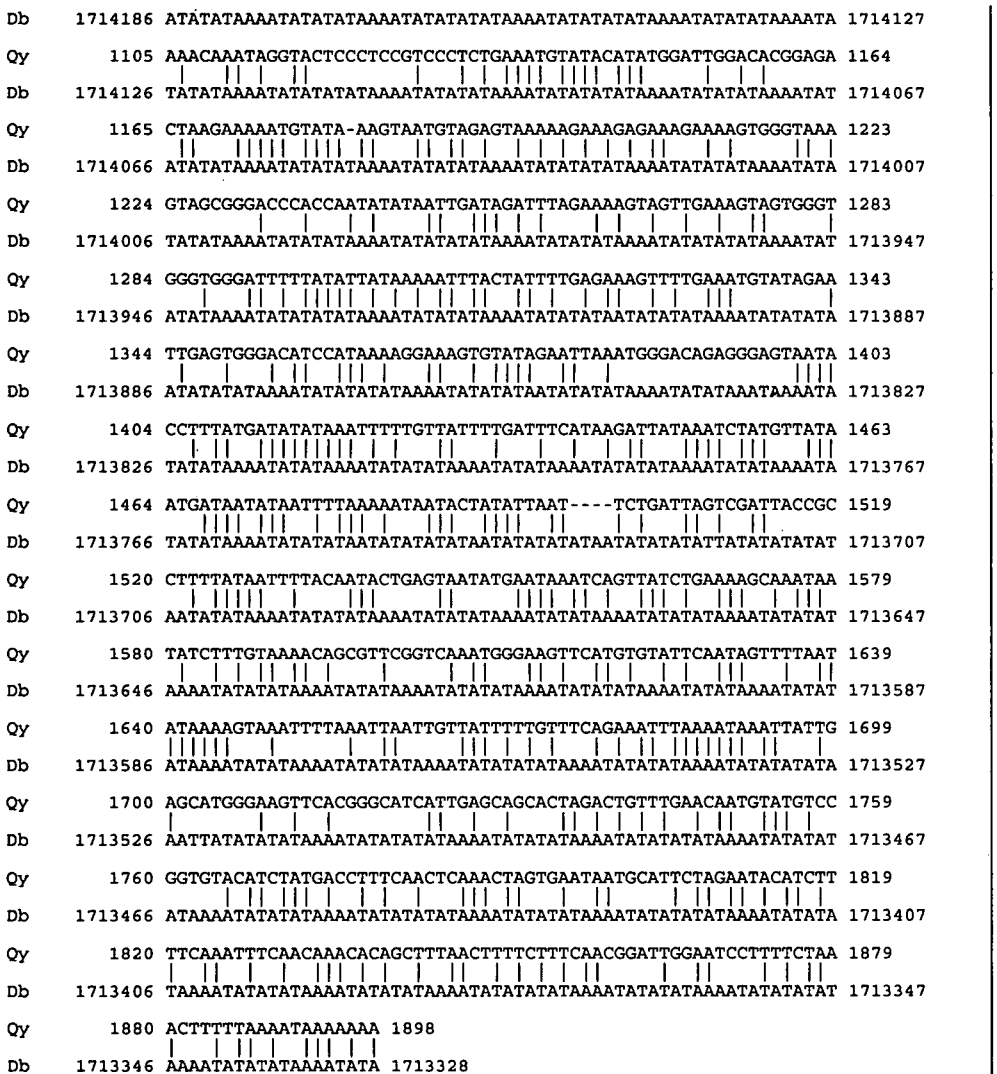
```
Qy       548 ATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGTAAAATTC 607
             ||   | | ||  | | ||| | || ||| | |||    |    |    || ||||||
Db      1975 ATTAATATTAATTTTTATTTAAAATTTAATAAAATTTTAAATATATTTAAAATAAAATAT 2034

Qy       608 CCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTT 667
             |||||  | ||| |   | || |   ||  | | | || |   ||    | |||  |
Db      2035 ATTTTTAAATATTTTTTATTTTTAAAAAATTTTATTAAAAATATTTTAAATATTATAAAT 2094

Qy       668 TTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCAC 727
                 | | ||   | |||    ||       |  | |    || | |   |
Db      2095 AATAATTATAAATTATTTAATATAATTTTAATAATTTTTTATTTTATTAAAAATAAAAAA 2154
```

```
Qy       728 TGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAG 787
Db      2155 TAAAAAAAAAAAAAATTATAATTTTATTAAAAATTTAAAAAAATGAAAAAAAAAATAAAA 2214

Qy       788 CTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAA 847
Db      2215 AAAAAAAAAAATTTAATTTTAATAAAAATTTATTTTTTAATATTTAAAATAAATTAAAAA 2274

Qy       848 GAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTC------TTTAAAAGTAATGT 901
Db      2275 AATTTTAATTTTTTAATTTATTTATTTTTAAAATAATTTATTAATTTTTGTATTTAAATA 2334

Qy       902 AAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATAT 961
Db      2335 AATTTTATTTAATTTTTAAATATAATTTTTTATTTTTATAAATAATTTATTAAATTTTAT 2394

Qy       962 TAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTACT 1021
Db      2395 TATTTTTTGAAAATTAAAAAAAAAAATAATTTATTTATTTTTAAAAATTAATT---TTTTT 2451

Qy      1022 GATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAA 1081
Db      2452 TTATTTTTATTTATATATTTTTTATTATTTTGTTTTAAAATAATTAAAATATTAATGAAA 2511

Qy      1082 AGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTA 1141
Db      2512 AATAAAAAAAAATAAATTAAATTATAAAATAAATTAAAAAATTTAAAAATTTAAATAAAA 2571

Qy      1142 TACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGA 1201
Db      2572 TTTTTAAAAAATAAAAAATAAATGAAAAAAAAAATTATAAATATAAATAATTTAAAAAAA 2631

Qy      1202 AAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGA 1261
Db      2632 AAAAATAAAAAAAAATGTTAAAATTTAAAATTTAAAAATAATAAATTTTTAAATTTTTT 2691

Qy      1262 AAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGA 1321
Db      2692 TAAATTTATTATTTTTTTTTTTTAATAAAAATAATTTGTTAAATAAAT----TATTAATA 2747

Qy      1322 GAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATT 1381
Db      2748 AAAATAAAAAAAATATTTAATTGTTTATTTTTAATAAATTTAAATGAATTTTGTTTATTA 2807

Qy      1382 AAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCAT 1441
Db      2808 AAATTGTTATTAATAATAAATGATATTTT--TTTATAAAATATTAAAAATTTGAAATTAT 2865

Qy      1442 AAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATTC 1501
Db      2866 TTTATGAAAAAAATTTTGTTGAAAAAAATAAAAAATTTGAAAATGAAAAAAAAATGTTAA 2925

Qy      1502 TGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGT 1561
Db      2926 AATTTTATGAAATGTAAAATTTAAATGAATTATTATTGTTAAAAAAATAAAAAAAAATTA 2985

Qy      1562 TATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGT 1621
Db      2986 TTGAAAAATTGATATATTTTAAAATGATGATGTTTAAATTTTATAAATATAAATTAAAAA 3045

Qy      1622 GTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAA 1681
Db      3046 TAAAATAAATGTTATAATTGTGAATGAAATTTGAATTATTGTTATTTTTATGTGTAATTT 3105

Qy      1682 ATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGAGCAGCACTAGACTG 1741
Db      3106 AATAAAAAATTAATTTTTTTAAAATGAATTTAATGATGATTGAAAAATTGTTTTATATAT 3165

Qy      1742 ---TTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATA 1798
Db      3166 AATTTTAAATAATAATATTTTTAATAAAAATTTTGTTGTTTGAATTGTAAATAAAATTTT 3225

Qy      1799 ATGCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAA 1858
Db      3226 AATTTATTATATTTTTTTTTAATAATTTGATTATTATTTTTAATTAATTGATTGTAAAAA 3285

Qy      1859 CGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATATTTA 1918
Db      3286 AATTTTTGTTATTGTTGTTGTTATTTTTTTGAAAAAAAAAAAATATAATTGTAATGTAAA 3345

Qy      1919 TCAACACCTCAACATTGATGTTA 1941
Db      3346 AAAATGAATAAAATTTGAAATGA 3368
```

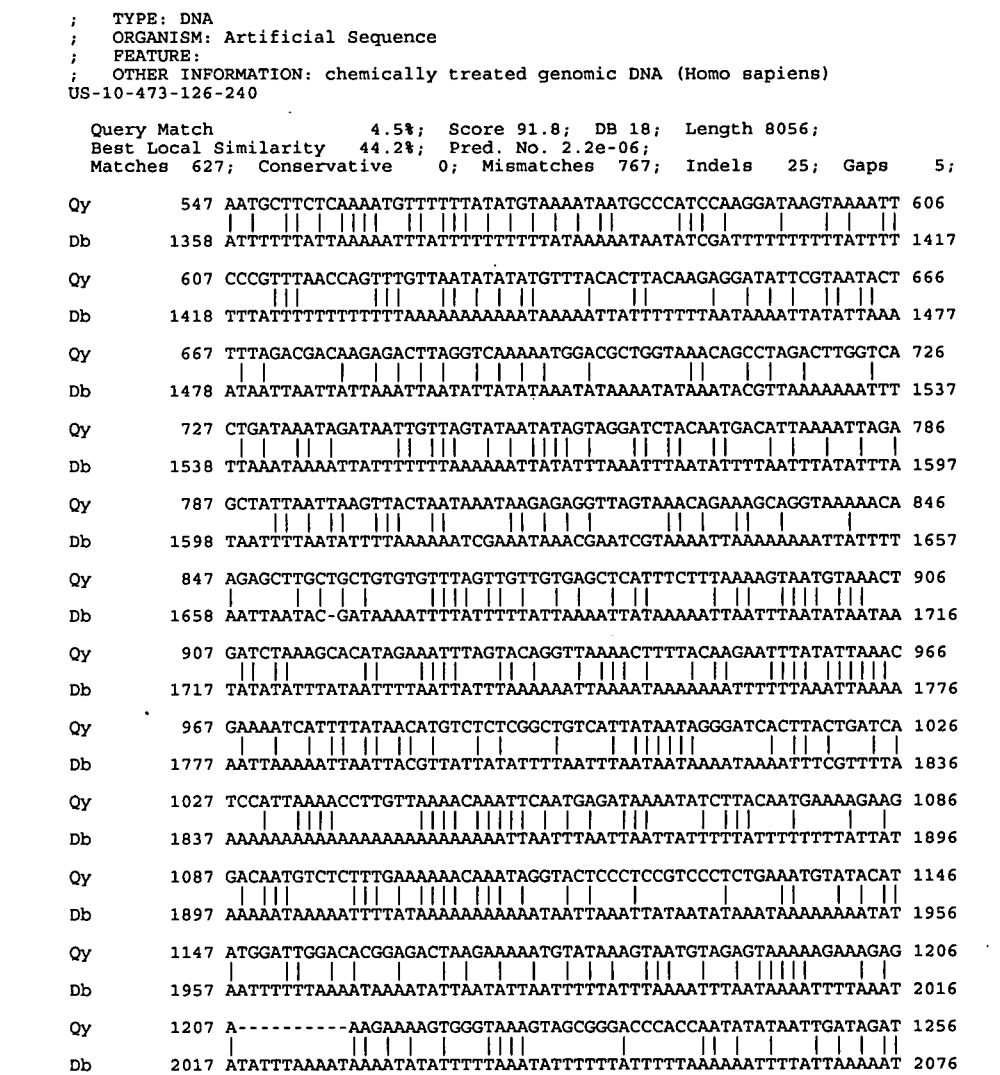
RESULT 2
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 4.5%; Score 93; DB 15; Length 3673778;
Best Local Similarity 43.8%; Pred. No. 1.1e-05;
Matches 587; Conservative 0; Mismatches 745; Indels 7; Gaps 4;

```
Qy        566 TTTATATGTAAAATAATGCCCATCCAAGGATAAGTAAAATTCCCGTTTAACCAGTTTGTT 625
Db    1714665 TATATATATAAAATATATATATAAAATATATATATAAAATATATATATAAAATATATA 1714606

Qy        626 AATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAAGAGACTT 685
Db    1714605 TATATAAAATATATACATAAAATATATATATATAAAATATATACATAAAATATATATAAA 1714546

Qy        686 AGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGATAATTGT 745
Db    1714545 ATATATATATAAAAAATATATATAAAATATATATAAAAAATATATATAAAATATATATAA 1714486

Qy        746 TAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGTTACTA 805
Db    1714485 AAAATATATATAAAATATATATAAAATATATATAAAATATATATAAAATATATATAAA 1714426

Qy        806 ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAGAGCTTGCTGCTGTGTGT 865
Db    1714425 ATATATATAAAATATATATATAAAATATATATAAAATATATATATAAAATATATATATAT 1714366

Qy        866 TTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGATCTAAAGCACATAGAAA 925
Db    1714365 AAAATATATATATAAAATATATATATAAAATATATAAAATATATATAAAATATATATAAA 1714306

Qy        926 TTTAGTACAGGTTAAAACTTTTACAAGAATTTATATTAAACGAAAATCATTTTATAACAT 985
Db    1714305 ATATATAAAATATATATATAAAATATATATATAAAATATA-TAAAATATATATAAAATAT 1714247

Qy        986 GTCTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCATCCATTAAAACCTTGTTAA 1045
Db    1714246 ATATAAAATATATAAAATATATATAAAATATATATATAAAATATATAAAATATATATAAA 1714187

Qy       1046 AACAAATTCAATGAGATAAAATATCT-TACAATGAAAAGAAGGACAATGTCTCTTTGAAA 1104
```

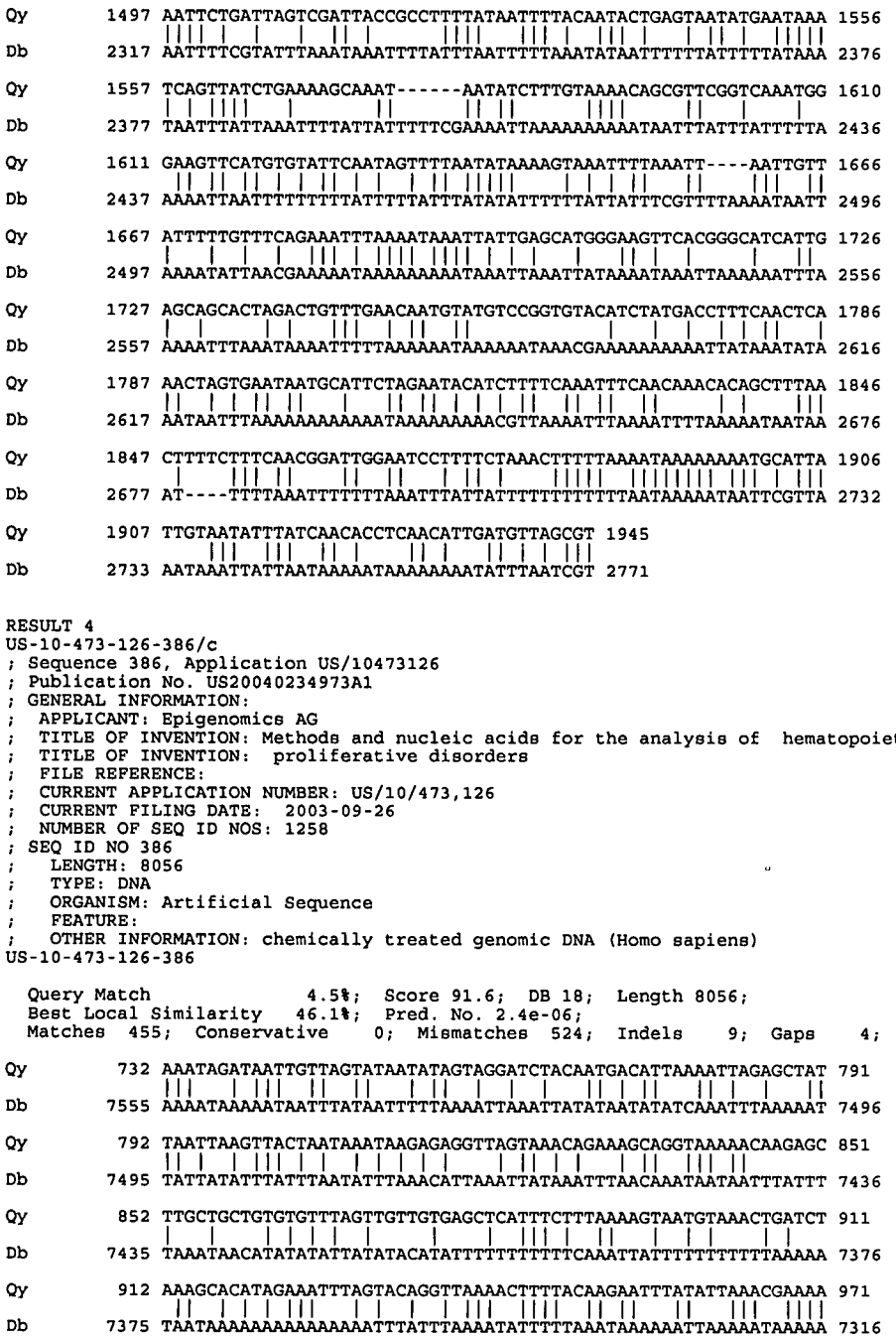


RESULT 3
US-10-473-126-240
; Sequence 240, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 240
; LENGTH: 8056



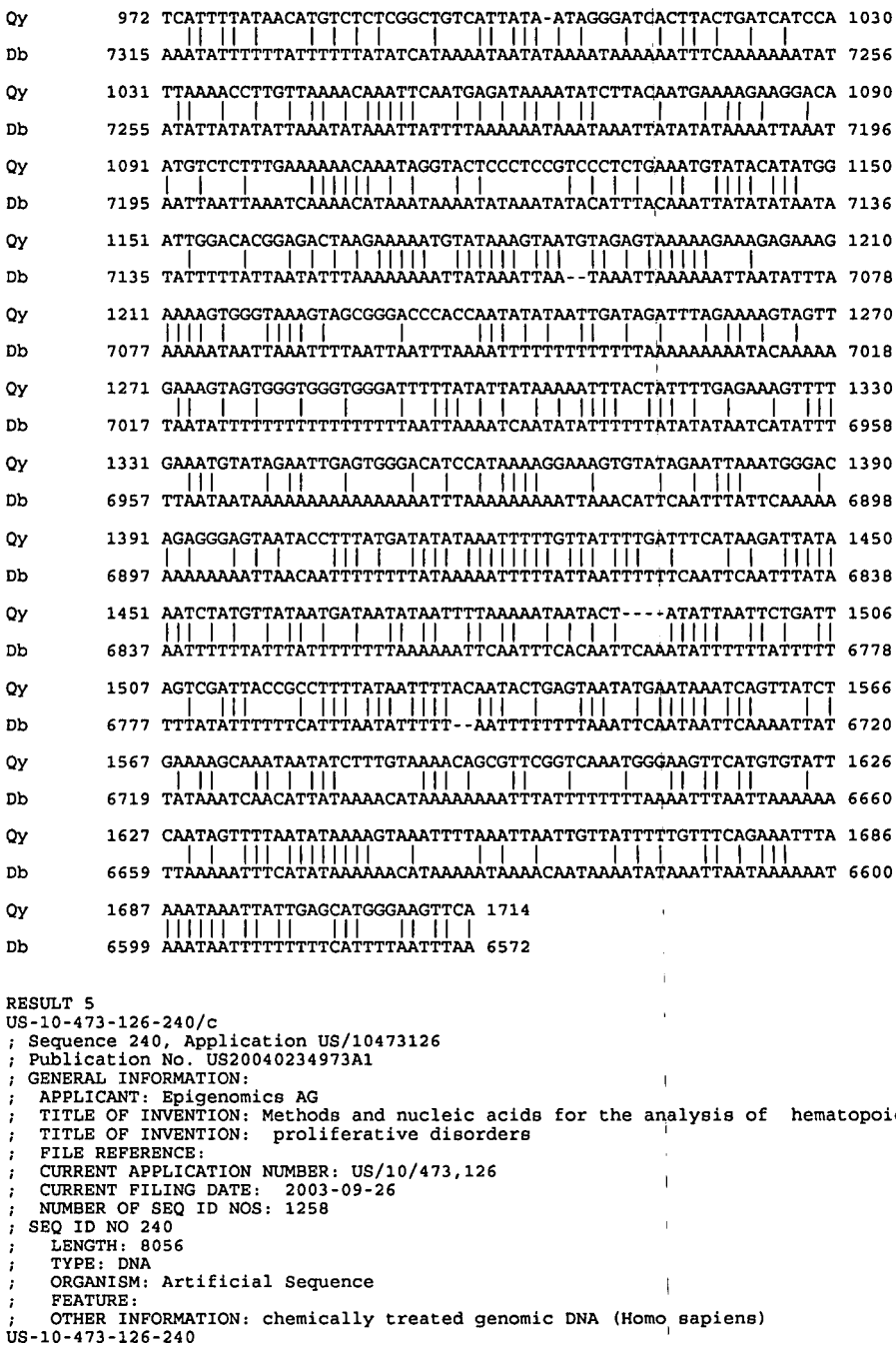
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-240

Query Match 4.5%; Score 91.8; DB 18; Length 8056;
Best Local Similarity 44.2%; Pred. No. 2.2e-06;
Matches 627; Conservative 0; Mismatches 767; Indels 25; Gaps 5;



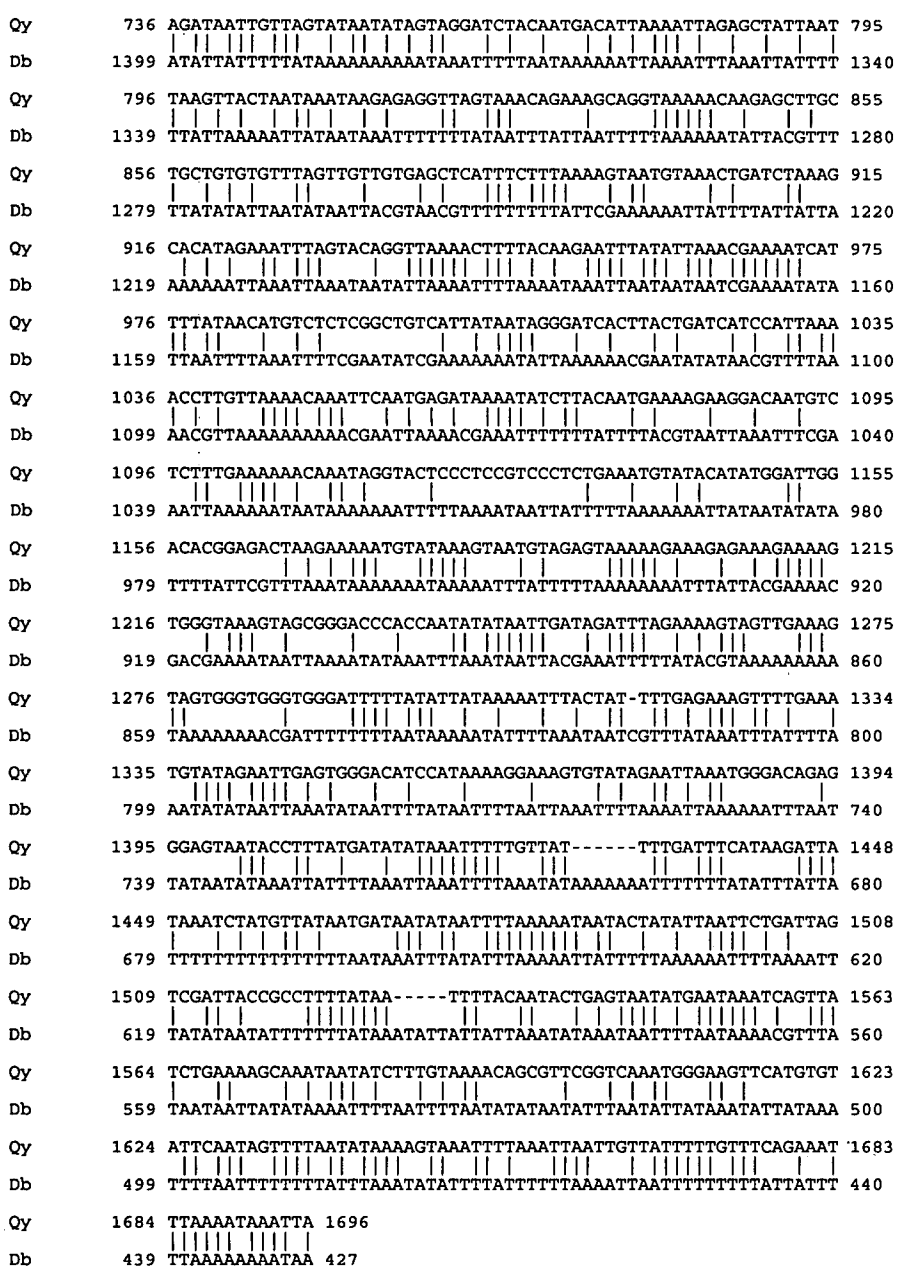
RESULT 4
US-10-473-126-386/c
; Sequence 386, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 386
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-386

Query Match 4.5%; Score 91.6; DB 18; Length 8056;
Best Local Similarity 46.1%; Pred. No. 2.4e-06;
Matches 455; Conservative 0; Mismatches 524; Indels 9; Gaps 4;



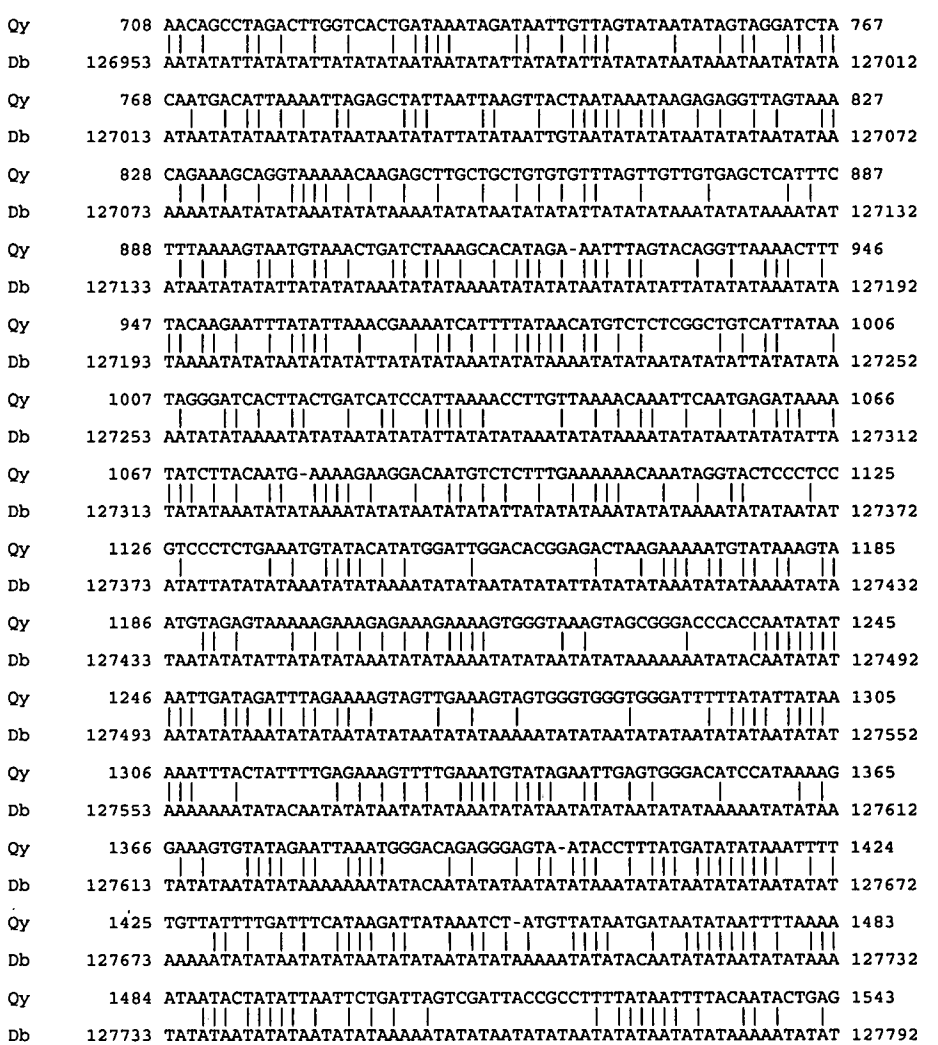
RESULT 5
US-10-473-126-240/c
; Sequence 240, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 240
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-240

Query Match 4.3%; Score 88.6; DB 18; Length 8056;
Best Local Similarity 45.4%; Pred. No. 8.2e-06;
Matches 442; Conservative 0; Mismatches 519; Indels 12; Gaps 3;



RESULT 6
US-10-211-179-11
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOTYROSYL PHOSPHATASE ACTIVATOR EXP.
; FILE REFERENCE: PTS-0011
; CURRENT APPLICATION NUMBER: US/10/211,179
; CURRENT FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 119
; OTHER INFORMATION: n = a, t, c, or g
US-10-211-179-11

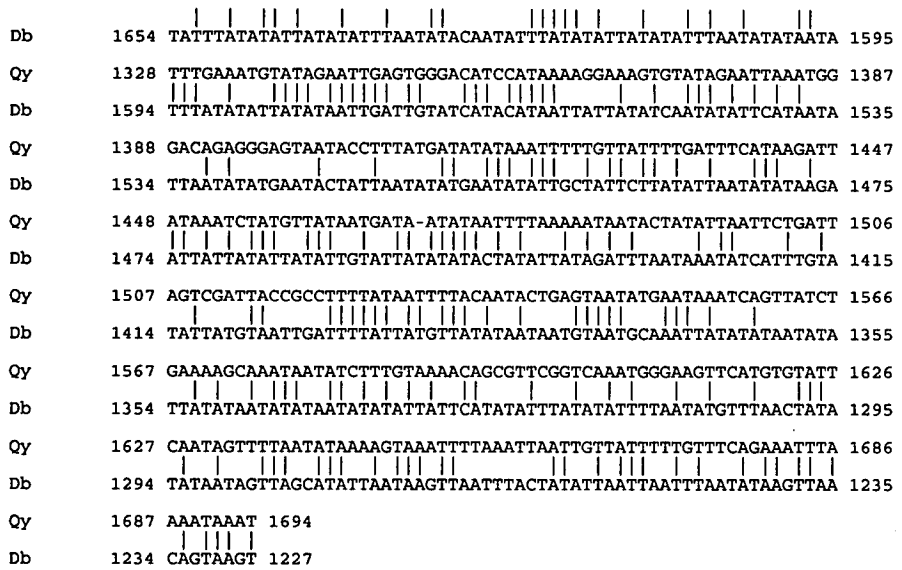
Query Match 4.0%; Score 83; DB 16; Length 158001;
Best Local Similarity 46.0%; Pred. No. 0.00023;
Matches 458; Conservative 0; Mismatches 530; Indels 7; Gaps 5;



```
Qy       1544 TAATATGAATAAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGT 1603
Db     127793 AATATATAATATATAATATAT---AAAAATATATAATATATAATATATAATATATAAAAA 127849

Qy       1604 CAAATGGGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATT 1663
Db     127850 TATATATAGCATATAAAAATATATTATACATTATATATAAAAATATATTATATATAATAT 127909

Qy       1664 GTTATTTTTGTTTCAGAAATTTAAAATAAATTATT 1698
Db     127910 ATTATATATATTATATATATATATATAATTTTTTT 127944


RESULT 7
US-10-240-453-206/c
; Sequence 206, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
;  TITLE OF INVENTION:  Transcription
;  TITLE OF INVENTION:  by Means of Assessing the Methylation Status of Genes Associated
;  TITLE OF INVENTION:  With DNA TRanscription
;  FILE REFERENCE: 5013.1009
;  CURRENT APPLICATION NUMBER: US/10/240,453
;  CURRENT FILING DATE:  2002-10-02
;  PRIOR APPLICATION NUMBER: PCT/EP01/03973
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 206
;   LENGTH: 11745
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (9105)
US-10-240-453-206

  Query Match             3.7%;  Score 75;  DB 15;  Length 11745;
  Best Local Similarity   42.7%;  Pred. No. 0.0025;
  Matches  499;  Conservative    0;  Mismatches  665;  Indels    5;  Gaps    2;

Qy        768 CAATGACATTAAAATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAA 827
Db       4805 CGAATAAATAATAATTCTATTTTCTAAATAAAAAAAATTATAATAAAAAAAACTAAATTT 4746

Qy        828 CAGAAAGCAGGTAAAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTC 887
Db       4745 AATAAATAAAAATAAAATCTAAAATAATATCAATAACTCTATTTTAACCGTATCAAAATA 4686

Qy        888 TTTAAAAGTAATGTAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTT 947
Db       4685 CTCATTACACATCTACCTAAAAATATTAAATAACCACTATATATATAAATCTAAAACTTA 4626

Qy        948 ACAAGAATTTATATTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAAT 1007
Db       4625 AAAAAAAAATAAACTAAAAAACCATTTTTAAAATACATCACAATAAAAATAATATTTAAA 4566

Qy       1008 AGGGATCACTTACTGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAAT 1067
Db       4565 ACCTTAAAATTAACTCGAATCACCTAAAA---AAAAAAAACAAATAAAAAAAAAAAAAAA 4509

Qy       1068 ATCTTACAATGAAAAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGT 1127
Db       4508 ACCAAAAACAAAAAACTAAATCACTCTAACATTTAAAAATCTAATAAAAAACAAAAAACA 4449

Qy       1128 CCCTCTGAAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAAT 1187
Db       4448 ACAAAAAAAAATAAAAAATACCCATTAAAATAAAAAAAAAACTATACAAATAAATATCAAA 4389

Qy       1188 GTAGAGTAAAAAGAAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAA 1247
Db       4388 ATAACATAAAAAAAAAAAACCTTTCAAAACAATAAAAATA--ATCAATTATCATATATACT 4331

Qy       1248 TTGATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAA 1307
Db       4330 TTTAAATCCTAAAAACAAAACTAAAAATCATTTAATTTTATCAATATATAACTCAATAAT 4271

Qy       1308 ATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGA 1367
Db       4270 TATTTTAATTAATTAAAACTCTTTTATTAAAATTAAAAAAATAATATCCTATTAAAATAA 4211

Qy       1368 AAGTGTATAGAATTAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGT 1427
Db       4210 ATTAAAAAAATAAAAAATAATAAAATAAAAACAACAAATATAAACAACAATTATAAAAAA 4151

Qy       1428 TATTTTGATTTCATAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAA 1487
Db       4150 TTCTACTAATAAAAAAAACAAAAAAATAAACGTTATATAAAATATAAAATCAAAAAAAAA 4091

Qy       1488 TACTATATTAATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAAT 1547
Db       4090 TTTTCTTTAAATAAATAATATTAAAACAACTCTATATATCAATAAAAATAACCCAAATAA 4031

Qy       1548 ATGAATAAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAA 1607
Db       4030 AAAAAAAACACAAATATTACAAAACAAAAAAAAATACTTACAAAAATAAAATCTTTAAAA 3971

Qy       1608 TGGGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTA 1667
Db       3970 AAACAATTAAAAATAAAATTAAAAACACAAATAAAAAAACTAACCTTAAATAAAAAAAAA 3911

Qy       1668 TTTTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGA 1727
Db       3910 AAAACTTTCCATTTTAACAAAAAAAAAAATATATTATATATAAATTTAAAAACAAATTAC 3851

Qy       1728 GCAGCACTAGACTGTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAA 1787
Db       3850 TAAAATTAATAATAACAAAACAAAATAATTCTTATCCATTTAAATTTATTTTCTCTTAAA 3791

Qy       1788 ACTAGTGAATAATGCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAAC 1847
Db       3790 AACATAAAACAAATTCATCAACCAAAAATAAAAAAAAAAATAAATATTAAAATTTTCAAAA 3731

Qy       1848 TTTTCTTTCAACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTAT 1907
Db       3730 TTAAAAAAAAAAATATTCAAAATAATCACCTAAAAAAAAAAAAAAATATATTTTTCTAAAA 3671

Qy       1908 TGTAATATTTATCAACACCTCAACATTGA 1936
Db       3670 TTAAATAATAAAACTACTAACAAAATTAA 3642


RESULT 8
US-10-311-455-490/c
; Sequence 490, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;  TITLE OF INVENTION:  cytosine methylation
```

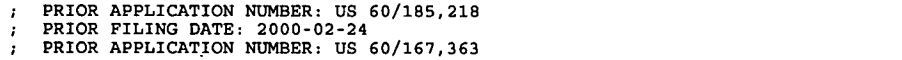
```
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 490
;   LENGTH: 5930
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-490

  Query Match             3.5%;  Score 71;  DB 15;  Length 5930;
  Best Local Similarity   46.4%;  Pred. No. 0.01;
  Matches  371;  Conservative    0;  Mismatches  420;  Indels    8;  Gaps    4;

Qy        897 AATGTAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATT 956
              ||| ||| |  ||| |  |||  ||| || ||| |||| || | |
Db       5154 AATTAAAAAATAAAAAAAATAATATAAAAACATTTACTACTTTAAATATTTATAAACAAT 5095

Qy        957 TATATTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCAC 1016
              | ||||| || | | | ||| | |||  || | |  |  || |
Db       5094 ATTTTTAAAAAAATTTACTCATCAAACTTATCTAAAACCTCAAACTTCCCAATTTATTAT 5035

Qy       1017 TTACTGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAA 1076
              || |   |   || ||| |   | |||  ||  |||  | | |||||   || ||
Db       5034 CTAATTTATAATATTTCAAAACAATATTAAAATAAAACAAAAATACAAAATTAAATAAAA 4975

Qy       1077 TGAAAAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAA 1136
              |||| ||    | |  || | | ||||||    ||   |      | |   ||
Db       4974 AAAAAACAATATAATTTACTTTATAAAAAAATTCTTAATAAAATAAAAAAAACTTAAAAA 4915

Qy       1137 ATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAA 1196
              || ||| |  | | | | |  | | || ||||||  | ||  ||   | |||
Db       4914 CCATAAACAAAAATAATAAATAAT-ATAAAAAAAAAAATAACTTAAAATATACCAATTAA 4856

Qy       1197 AAAGAAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGA--TAG 1254
              || |||    |    |   ||| ||   |  ||||||||||||  |  ||
Db       4855 TAATAAAACTCTAACTTAAAAATTAATATATAAAAATATACCAATATATAATATATTTAT 4796

Qy       1255 ATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACT 1314
              | |  |||| | |  |||         |   |  |||  |||||  ||| |  ||| |
Db       4795 AATATAAAAATAAATATAAAAACTAATATAATTAATATTAATATATATTAATATATTAAT 4736

Qy       1315 ATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTA 1374
              || |  |  ||  ||  || | || |   |  |  |   ||| |   ||
Db       4735 ATATAAATTAACATTATAATTATACATTCGTATATAAAATTATTATATATATTAATAAAT 4676

Qy       1375 TAGAATTAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTG 1434
              ||  | ||||    | | |      || ||| | ||  |||||  ||| |  | | |
Db       4675 TAATAATAAAACTAATAAAATTTATATTACTTATAAAATATAAAAAATACTATAAAAT-- 4618

Qy       1435 ATTTCATAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATA 1494
              || | ||    ||  || ||| || |||  ||   |  ||| |||   ||  |||
Db       4617 CTTCCCTATATATAATTATATATAATAAAATATTATATTCTATTTCTAAACCCTAACATA 4558

Qy       1495 TTAATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATA 1554
              ||||  |   |  | ||      | |||  | |   |    || || |
Db       4557 AAAATTTAAAAACTTATAATAATCTTAAACAAAATAAAATTAATAAAAACTTAATTAACA 4498

Qy       1555 AATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAG 1614
                | |  |||  ||||| |  ||| || || |  ||  |  ||      |
Db       4497 TTCATAAAAAATCTAAAAAAAATATTTATTTTCTATAATAAAATTATATTCAACCTTCAA 4438

Qy       1615 TTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGT 1674
              |  |   |   ||  | |||| |  |  || ||||  |||||      ||||
Db       4437 ATATTAAAAAAAACATCTCTCAAATAAATTAAAAACTTTT---TTAATAAAATATTTTAA 4381

Qy       1675 TTCAGAAATTTAAAATAAA 1693
              |||| ||||||||   |||
Db       4380 TTCAAAAATTTAATTCAAA 4362


RESULT 9
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
;  APPLICANT: Epigenomics AG
;  TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MH
;  FILE REFERENCE: E01/1208/WO
;  CURRENT APPLICATION NUMBER: US/10/312,841
;  CURRENT FILING DATE:  2002-12-30
;  NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
;   LENGTH: 3673778
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (3294164)
US-10-312-841-1

  Query Match             3.4%;  Score 70.6;  DB 15;  Length 3673778;
  Best Local Similarity   42.2%;  Pred. No. 0.11;
  Matches  644;  Conservative    0;  Mismatches  874;  Indels    9;  Gaps    4;

Qy        149 TCTTTTCCAAACACTTTATTAACTTTTTTACTTCTCATTTCTACTCCACTTCTTTGCTAT 208
              | ||||  | | | |||||  |  | |||  |    ||||    |  |  | ||   |||
Db    1713913 TATTTTATATATATTTTATATATATATTTTATATATATTTTATATATATATTTTATATAT 1713972

Qy        209 AAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAAGATCATTCATAA 268
              |    | | ||   || | |  |  |  |      | |      |  ||| |   |   |
Db    1713973 ATTTTATATATATATTTTATATATATTTTATATATATATTTTATATATATTTTATATATA 1714032

Qy        269 ATGTATCTTTCAATTTTAGGATAACAATACGTGAACAGGGTTATTTTTTAACGTGTCAAC 328
              | | | |  ||||||      |   ||  |  | |   |  | | |    | |  |
Db    1714033 TATTTTATATATATTTTATATATATTTTATATATATATTTTATATATATTTTATATATAT 1714092

Qy        329 AAATTCTAATAATTTTACCTGGCCGGTGAACACCGTCTTCCAAGATAATATATTTTAATT 388
              |  || ||   ||||||  |      |  | |    ||   | ||| | ||| |  ||
Db    1714093 ATTTTATATATATTTTATATATATATTTTATATATATTTTATATATATTTTATATATATA 1714152

Qy        389 TTGTAGCCTCCCTTTTAACCAAATTCGCATGCAGGACGACTTAGGTGAATACACATTGTA 448
              || ||       ||| |   |  |   ||  |        |    | ||| |  || ||
Db    1714153 TTTTATATATATATTTTATATATATTTTATATATTTTATATATATTTTATATATTTTATA 1714212

Qy        449 CTGTGAGTCTTTAAACAAAGAACAAGTGGTTCATGCTCAGCCATCAAAATTGACAAAACC 508
              |  | | || | |    | |  |  |  ||           | | || | | |
Db    1714213 TATATATTTTATATATATTTTATATATTTTATATATATTTTATATATATTTTATATATT 1714272

Qy        509 CGACACAACACTCTATCCACGTACTATACTTTTGGCCGAATGCTTCTCAAAATGTTTTTT 568
              |  | | |||  |  || | ||  | |      |   || |  | || |||| |
Db    1714273 TATATATATATTTTATATATATATTTTATATATTTTATATATATTTTATATATATTTTAT 1714332

Qy        569 ATATGTAAAATAATGCCCATCCAAGGATAAGTAAAAT---TCCCGTTTAACCAGTTTGTT 625
              |||| | | |||         |  |  |||  | | ||   |   ||||   |  | ||
Db    1714333 ATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATATATATATTT 1714392

Qy        626 AATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAAGAGACTT 685
              ||||||||  ||    |        ||||   | |||  | |    | |     ||
Db    1714393 TATATATATTTTATATATATATTTTATATATATTTTATATATATTTTATATATATATTTT 1714452

Qy        686 AGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGATAATTGT 745
              |  |  |  |  |      | |       |   |  |  | || ||| ||  ||
Db    1714453 ATATATATTTTATATATATTTTATATATATTTTTTATATATATTTTATATATATTTTTA 1714512

Qy        746 TAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGTTACTA 805
```

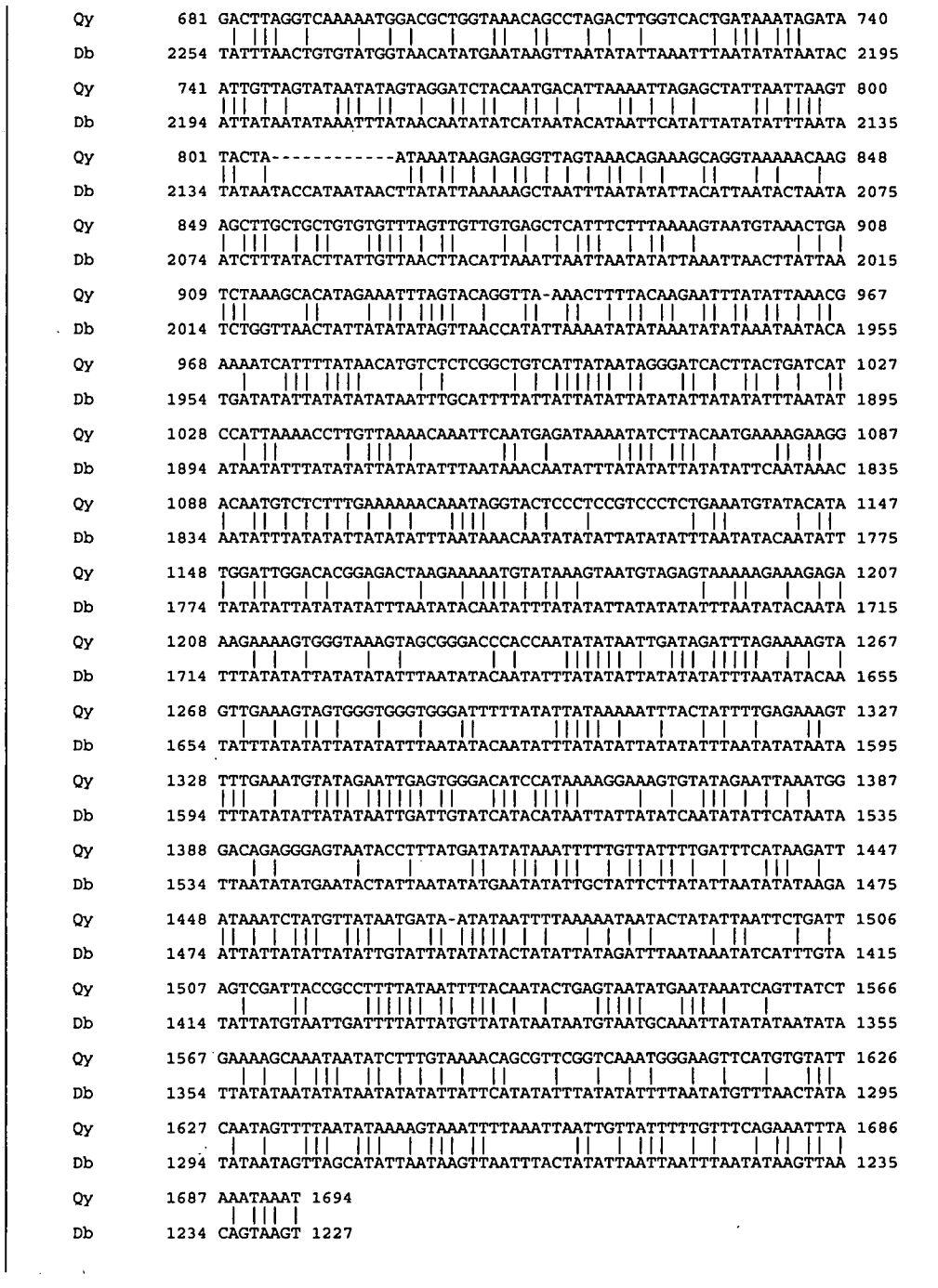
```
Db    1714513 TATATATTTTATATATATTTTTTATATATATATTTTATATATATTTTATGTATATATTTT 1714572

Qy        806 ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAGAGCTTGCTGCTGTGTGT 865

Db    1714573 ATATATATATATTTTATGTATATATTTTATATATATATATTTTATATATATATTTTATAT 1714632

Qy        866 TTAGTTGTTGT-GAGCTCATTTCTTTAAAAGTAATGTAAACTGATCTAAAGCACATAGAA 924

Db    1714633 ATATATATTTTATATATATATTTTATATATATATTTTATATATTTTATATGTATATTT 1714692

Qy        925 ATTTAGTACAGGTTAAAACTTTTACAAGAATTTATATTAAACGAAAATCATTTTATAACA 984

Db    1714693 TATATGTATATTTTATATGTATATTTTATATGTATATTTTATATGTATATTTTATATATA 1714752

Qy        985 TGTCTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCATCCATTAAAACCTTGTTA 1044

Db    1714753 TATTTTATATGTATATTTTATATATATATTTTATATATATTTTATATATATATTTTATAT 1714812

Qy       1045 AAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGGACAATGTCTCTTTGAAA 1104

Db    1714813 ATATATTATTTATATTTTATATATATAGTATATATATTTTATATGTATATATTTTATATA 1714872

Qy       1105 AAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATATGGATTGGACACGGAGA 1164

Db    1714873 TATTTTATATATATATATTTTATATATATATTTTATATATATTTTATATATATATTTTGT 1714932

Qy       1165 CTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGAAAGAAAAGTGGGTAAAG 1224

Db    1714933 ATATATATATTTTGTATATATATATTTTATATATATATTTTATATATATATATTTTAT 1714992

Qy       1225 TAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGTG 1284

Db    1714993 ATATATATATATTTTATATATATTTTATATATATATA----TATTTTATATATATATTTT 1715048

Qy       1285 GGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAAT 1344

Db    1715049 ATATATATATATATTTTATATATATATTTTATATATATATATTTTATATATATATTT 1715108

Qy       1345 TGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAATAC 1404

Db    1715109 TTATATATATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATATAT 1715168

Qy       1405 CTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATTATAAATCTATGTTATAA 1464

Db    1715169 ATATTTTATATATATATATTTTATATATATATTTTATATATATATATTTTATATATAT 1715228

Qy       1465 TGATAATATAATTTTAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGCCTTTT 1524

Db    1715229 TTTATATATATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATATA 1715288

Qy       1525 ATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCTGAAAAGCAAATAATATCT 1584

Db    1715289 TATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATATAT-ATATAT 1715347

Qy       1585 TTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAATATAAA 1644

Db    1715348 TTTATATGTATATTTTATATATATATATTTTATATGTATATTTTATATATATATTTTATA 1715407

Qy       1645 AGTAAATTTTAAATTAATTGTTATTTT 1671

Db    1715408 TGTATATTTTATATGTATATTTTATAT 1715434


RESULT 10
US-10-027-632-113786/c
; Sequence 113786, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113786
;   LENGTH: 3252
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-113786

  Query Match             3.4%;  Score 70;  DB 13;  Length 3252;
  Best Local Similarity   43.6%;  Pred. No. 0.012;
  Matches  474;  Conservative    0;  Mismatches  600;  Indels   14;  Gaps    3;

Qy        621 TTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAAGA 680

Db       2314 TTAATACTAACAATTTTTATACTTATATTAATGAGTTTAATATGTTTATACTGTAATATA 2255

Qy        681 GACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGATA 740

Db       2254 TATTTAACTGTGTATGGTAACATATGAATAAGTTAATATATTAAATTTAATATATAATAC 2195

Qy        741 ATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGT 800

Db       2194 ATTATAATATAAATTTATAACAATATATCATAATACATAATTCATATTATATATTTAATA 2135

Qy        801 TACTA----------ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAG 848

Db       2134 TATAATACCATAATAACTTATATTAAAAAGCTAATTTAATATATTACATTAATACTAATA 2075

Qy        849 AGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGA 908

Db       2074 ATCTTTATACTTATTGTTAACTTACATTAAATTAATTAATATATTAAATTAACTTATTAA 2015

Qy        909 TCTAAAGCACATAGAAATTTAGTACAGGTTA-AAACTTTTACAAGAATTTATATTAAACG 967

Db       2014 TCTGGTTAACTATTATATATAGTTAACCATATTAAAATATATAAATATATAAATAATACA 1955

Qy        968 AAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCAT 1027

Db       1954 TGATATATTATATATATAATTTGCATTTTATTATTATATTATATATTATATATTTAATAT 1895

Qy       1028 CCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGG 1087

Db       1894 ATAATATTTATATATTATATATTTAATAAACAATATTTATATATTATATATTCAATAAAC 1835

Qy       1088 ACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATA 1147

Db       1834 AATATTTATATATTATATATTTAATAAACAATATATTATATATTTAATATACAATATT 1775

Qy       1148 TGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGA 1207

Db       1774 TATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAATA 1715

Qy       1208 AAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTA 1267

Db       1714 TTTATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAA 1655

Qy       1268 GTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGT 1327
```



RESULT 11
US-10-027-632-113787/c
; Sequence 113787, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113787
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113787

Query Match 3.4%; Score 70; DB 13; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;





RESULT 12
US-10-027-632-113788/c
; Sequence 113788, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113788
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113788

Query Match 3.4%; Score 70; DB 13; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;

```
Qy       621 TTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAAGA 680
Db      2314 TTAATACTAACAATTTTTATACTTATATTAATGAGTTTAATATGTTTATACTGTAATATA 2255

Qy       681 GACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGATA 740
Db      2254 TATTTAACTGTGTATGGTAACATATGAATAAGTTAATATATTAAATTTAATATATAATAC 2195

Qy       741 ATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGT 800
Db      2194 ATTATAATATAAATTTATAACAATATATCATAATACATAATTCATATTATATATTTAATA 2135

Qy       801 TACTA------------ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAG 848
Db      2134 TATAATACCATAATAACTTATATTAAAAAGCTAATTTAATATATTACATTAATACTAATA 2075

Qy       849 AGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGA 908
Db      2074 ATCTTTATACTTATTGTTAACTTACATTAAATTAATTAATATATTAAATTAACTTATTAA 2015

Qy       909 TCTAAAGCACATAGAAATTTAGTACAGGTTA-AAACTTTTACAAGAATTTATATTAAACG 967
Db      2014 TCTGGTTAACTATTATATATAGTTAACCATATTAAAATATATAAATATATAAATAATACA 1955

Qy       968 AAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCAT 1027
Db      1954 TGATATATTATATATATAATTTGCATTTTATTATTATATTATATATTATATATTTAATAT 1895

Qy      1028 CCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGG 1087
Db      1894 ATAATATTTATATATTATATATTTAATAAACAATATTTATATATTATATATTCAATAAAC 1835

Qy      1088 ACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATA 1147
Db      1834 AATATTTATATATTATATATTTAATAAACAATATATATTATATATTTAATATACAATATT 1775

Qy      1148 TGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGA 1207
Db      1774 TATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAATA 1715

Qy      1208 AAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTA 1267
Db      1714 TTTATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAA 1655

Qy      1268 GTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGT 1327
Db      1654 TATTTATATATTATATATTTAATATACAATATTTATATATTATATATTTAATATATAATA 1595

Qy      1328 TTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGG 1387
Db      1594 TTTATATATTATATAATTGATTGTATCATACATAATTATTATATCAATATATTCATAATA 1535

Qy      1388 GACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATT 1447
Db      1534 TTAATATATGAATACTATTAATATATGAATATATTGCTATTCTTATATTAATATATAAGA 1475

Qy      1448 ATAAATCTATGTTATAATGATA-ATATAATTTTAAAAATAATACTATATTAATTCTGATT 1506
Db      1474 ATTATTATATTATATTGTATTATATATACTATATTATAGATTTAATAAATATCATTTGTA 1415

Qy      1507 AGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCT 1566
Db      1414 TATTATGTAATTGATTTTATTATGTTATATAATAATGTAATGCAAATTATATATAATATA 1355

Qy      1567 GAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATT 1626
Db      1354 TTATATAATATATAATATATTATTCATATATTTATATATTTTAATATGTTTAACTATA 1295

Qy      1627 CAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTA 1686
Db      1294 TATAATAGTTAGCATATTAATAAGTTAATTTACTATATTAATTAATTTAATATAAGTTAA 1235

Qy      1687 AAATAAAT 1694
Db      1234 CAGTAAGT 1227
```

RESULT 13
US-10-027-632-113786/c
; Sequence 113786, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113786
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113786

Query Match 3.4%; Score 70; DB 15; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;

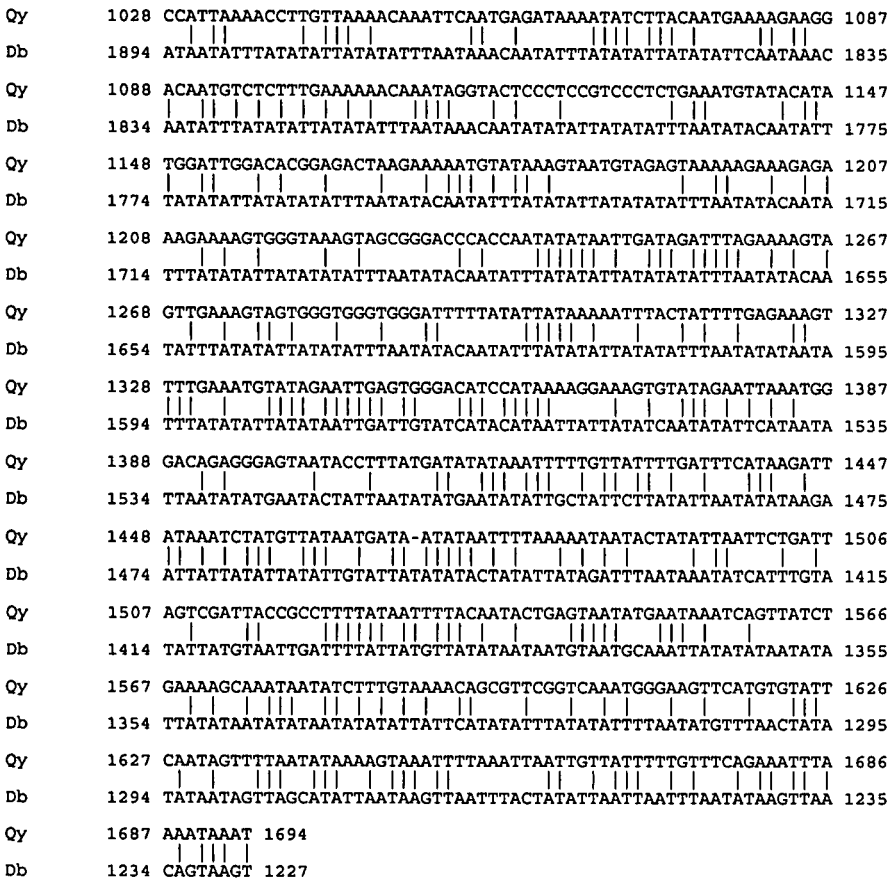
```
Qy       621 TTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAAGA 680
Db      2314 TTAATACTAACAATTTTTATACTTATATTAATGAGTTTAATATGTTTATACTGTAATATA 2255

Qy       681 GACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGATA 740
Db      2254 TATTTAACTGTGTATGGTAACATATGAATAAGTTAATATATTAAATTTAATATATAATAC 2195

Qy       741 ATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGT 800
Db      2194 ATTATAATATAAATTTATAACAATATATCATAATACATAATTCATATTATATATTTAATA 2135

Qy       801 TACTA------------ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAG 848
Db      2134 TATAATACCATAATAACTTATATTAAAAAGCTAATTTAATATATTACATTAATACTAATA 2075

Qy       849 AGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGA 908
Db      2074 ATCTTTATACTTATTGTTAACTTACATTAAATTAATTAATATATTAAATTAACTTATTAA 2015

Qy       909 TCTAAAGCACATAGAAATTTAGTACAGGTTA-AAACTTTTACAAGAATTTATATTAAACG 967
Db      2014 TCTGGTTAACTATTATATATAGTTAACCATATTAAAATATATAAATATATAAATAATACA 1955

Qy       968 AAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCAT 1027
Db      1954 TGATATATTATATATATAATTTGCATTTTATTATTATATTATATATTATATATTTAATAT 1895

Qy      1028 CCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGG 1087
Db      1894 ATAATATTTATATATTATATATTTAATAAACAATATTTATATATTATATATTCAATAAAC 1835

Qy      1088 ACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATA 1147
Db      1834 AATATTTATATATTATATATTTAATAAACAATATATATTATATATTTAATATACAATATT 1775

Qy      1148 TGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGA 1207
Db      1774 TATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAATA 1715

Qy      1208 AAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTA 1267
Db      1714 TTTATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAA 1655

Qy      1268 GTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGT 1327
Db      1654 TATTTATATATTATATATTTAATATACAATATTTATATATTATATATTTAATATATAATA 1595

Qy      1328 TTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGG 1387
Db      1594 TTTATATATTATATAATTGATTGTATCATACATAATTATTATATCAATATATTCATAATA 1535

Qy      1388 GACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATT 1447
Db      1534 TTAATATATGAATACTATTAATATATGAATATATTGCTATTCTTATATTAATATATAAGA 1475

Qy      1448 ATAAATCTATGTTATAATGATA-ATATAATTTTAAAAATAATACTATATTAATTCTGATT 1506
Db      1474 ATTATTATATTATATTGTATTATATATACTATATTATAGATTTAATAAATATCATTTGTA 1415

Qy      1507 AGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCT 1566
Db      1414 TATTATGTAATTGATTTTATTATGTTATATAATAATGTAATGCAAATTATATATAATATA 1355

Qy      1567 GAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATT 1626
Db      1354 TTATATAATATATAATATATATTATTCATATATTTATATATTTTAATATGTTTAACTATA 1295
```

```
Qy      1627 CAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTA 1686
Db      1294 TATAATAGTTAGCATATTAATAAGTTAATTTACTATATTAATTAATTTAATATAAGTTAA 1235

Qy      1687 AAATAAAT 1694
Db      1234 CAGTAAGT 1227
```

RESULT 14
US-10-027-632-113787/c
; Sequence 113787, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113787
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113787

Query Match 3.4%; Score 70; DB 15; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;

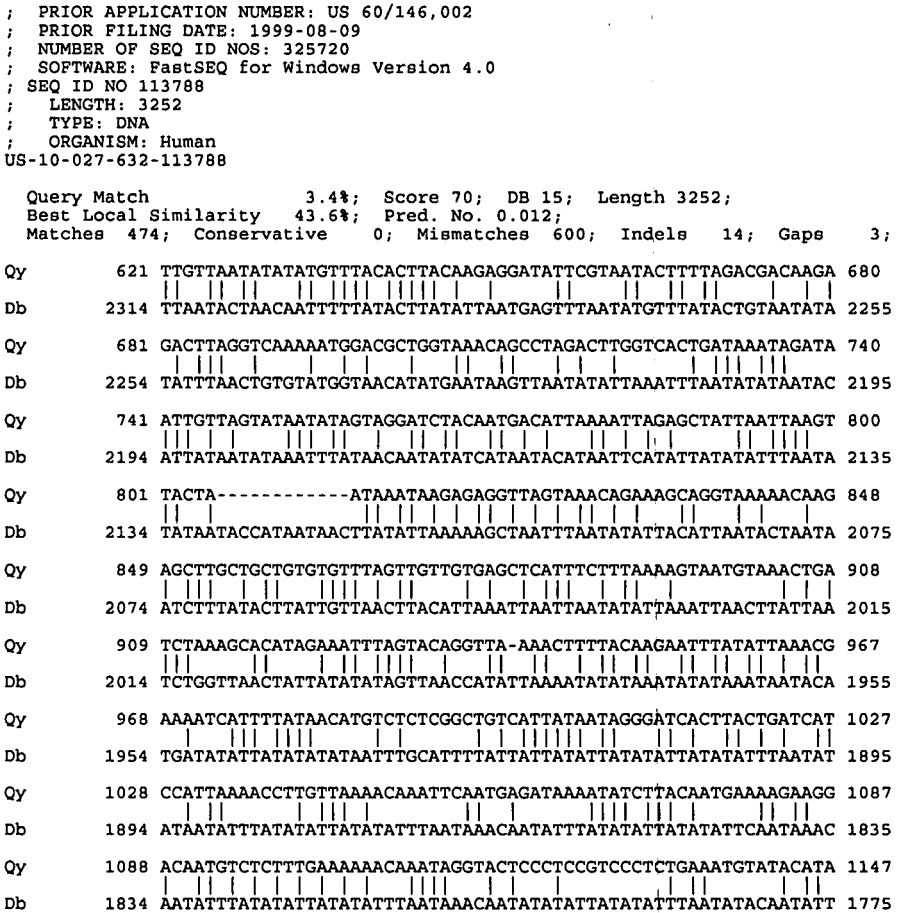
```
Qy       621 TTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAAGA 680
Db      2314 TTAATACTAACAATTTTTATACTTATATTAATGAGTTTAATATGTTTATACTGTAATATA 2255

Qy       681 GACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGATA 740
Db      2254 TATTTAACTGTGTATGGTAACATATGAATAAGTTAATATATTAAATTTAATATATAATAC 2195

Qy       741 ATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGT 800
Db      2194 ATTATAATATAAATTTATAACAATATATCATAATACATAATTCATATTATATATTTAATA 2135

Qy       801 TACTA------------ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAG 848
Db      2134 TATAATACCATAATAACTTATATTAAAAAGCTAATTTAATATATTACATTAATACTAATA 2075

Qy       849 AGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGA 908
Db      2074 ATCTTTATACTTATTGTTAACTTACATTAAATTAATTAATATATTAAATTAACTTATTAA 2015

Qy       909 TCTAAAGCACATAGAAATTTAGTACAGGTTA-AAACTTTTACAAGAATTTATATTAAACG 967
Db      2014 TCTGGTTAACTATTATATATAGTTAACCATATTAAAATATATAAATATATAAATAATACA 1955

Qy       968 AAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCAT 1027
Db      1954 TGATATATTATATATATAATTTGCATTTTATTATTATATTATATATTATATATTTAATAT 1895
```



```
Qy       1028 CCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGG 1087
Db       1894 ATAATATTTATATATTATATATTTAATAAACAATATTTATATATTATATATTCAATAAAC 1835

Qy       1088 ACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATA 1147
Db       1834 AATATTTATATATTATATATTTAATAAACAATATATATTATATATTTAATATACAATATT 1775

Qy       1148 TGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGA 1207
Db       1774 TATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAATA 1715

Qy       1208 AAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTA 1267
Db       1714 TTTATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAA 1655

Qy       1268 GTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGT 1327
Db       1654 TATTTATATATTATATATTTAATATACAATATTTATATATTATATATATTTAATATATAATA 1595

Qy       1328 TTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGG 1387
Db       1594 TTTATATATTATATAATTGATTGTATCATACATAATTATTATATCAATATATTCATAATA 1535

Qy       1388 GACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATT 1447
Db       1534 TTAATATATGAATACTATTAATATATGAATATATTGCTATTCTTATATTAATATATAAGA 1475

Qy       1448 ATAAATCTATGTTATAATGATA-ATATAATTTTAAAAATAATACTATATTAATTCTGATT 1506
Db       1474 ATTATTATATTATATTGTATTATATATACTATATTATAGATTTAATAAATATCATTTGTA 1415

Qy       1507 AGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCT 1566
Db       1414 TATTATGTAATTGATTTTATTATGTTATATAATAATGTAATGCAAATTATATATAATATA 1355

Qy       1567 GAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATT 1626
Db       1354 TTATATAATATATAATATATATTATTCATATATTTATATATTTTAATATGTTTAACTATA 1295

Qy       1627 CAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTA 1686
Db       1294 TATAATAGTTAGCATATTAATAAGTTAATTTACTATATTAATTAATTTAATATAAGTTAA 1235

Qy       1687 AAATAAAT 1694
Db       1234 CAGTAAGT 1227
```

RESULT 15
US-10-027-632-113788/c
; Sequence 113788, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113788
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113788

Query Match 3.4%; Score 70; DB 15; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;



```
Qy        621 TTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAAGA 680
Db       2314 TTAATACTAACAATTTTTATACTTATATTAATGAGTTTAATATGTTTATACTGTAATATA 2255

Qy        681 GACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGATA 740
Db       2254 TATTTAACTGTGTATGGTAACATATGAATAAGTTAATATATTAAATTTAATATATAATAC 2195

Qy        741 ATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGT 800
Db       2194 ATTATAATATAAATTTATAACAATATATCATAATACATAATTCATATTATATATTTAATA 2135

Qy        801 TACTA------------ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAG 848
Db       2134 TATAATACCATAATAACTTATATTAAAAAGCTAATTTAATATATTACATTAATACTAATA 2075

Qy        849 AGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGA 908
Db       2074 ATCTTTATACTTATTGTTAACTTACATTAAATTAATTAATATATTAAATTAACTTATTAA 2015

Qy        909 TCTAAAGCACATAGAAATTTAGTACAGGTTA-AAACTTTTACAAGAATTTATATTAAACG 967
Db       2014 TCTGGTTAACTATTATATATAGTTAACCATATTAAAATATATAAATATATAAATAATACA 1955

Qy        968 AAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCAT 1027
Db       1954 TGATATATTATATATATAATTTGCATTTTATTATTATATTATATATTATATATTTAATAT 1895

Qy       1028 CCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGG 1087
Db       1894 ATAATATTTATATATTATATATTTAATAAACAATATTTATATATTATATATTCAATAAAC 1835

Qy       1088 ACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATA 1147
Db       1834 AATATTTATATATTATATATTTAATAAACAATATATATTATATATTTAATATACAATATT 1775

Qy       1148 TGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGA 1207
Db       1774 TATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAATA 1715

Qy       1208 AAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTA 1267
Db       1714 TTTATATATTATATATATTTAATATACAATATTTATATATTATATATATTTAATATACAA 1655

Qy       1268 GTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGT 1327
Db       1654 TATTTATATATTATATATTTAATATACAATATTTATATATTATATATTTAATATATAATA 1595

Qy       1328 TTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGG 1387
Db       1594 TTTATATATTATATAATTGATTGTATCATACATAATTATTATATCAATATATTCATAATA 1535

Qy       1388 GACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATT 1447
Db       1534 TTAATATATGAATACTATTAATATATGAATATATTGCTATTCTTATATTAATATATAAGA 1475

Qy       1448 ATAAATCTATGTTATAATGATA-ATATAATTTTAAAAATAATACTATATTAATTCTGATT 1506
Db       1474 ATTATTATATTATATTGTATTATATATACTATATTATAGATTTAATAAATATCATTTGTA 1415
```

```
Qy       1507 AGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCT 1566
              |    ||     |||||| || ||| |  |    |||||   ||| |   |
Db       1414 TATTATGTAATTGATTTTATTATGTTATATAATAATGTAATGCAAATTATATATAATATA 1355

Qy       1567 GAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATT 1626
              | |  | |||  || | | | |  ||     |   | | |  | |   |   |||
Db       1354 TTATATAATATATAATATATTATTCATATATTTATATATTTTAATATGTTTAACTATA 1295

Qy       1627 CAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTA 1686
              |  |   |||  |||  | ||| ||         ||  | |||  | |   | || || |
Db       1294 TATAATAGTTAGCATATTAATAAGTTAATTTACTATATTAATTAATTTAATATAAGTTAA 1235

Qy       1687 AAATAAAT 1694
              | ||| |
Db       1234 CAGTAAGT 1227
```

Search completed: December 7, 2004, 06:11:14
Job time : 1222.45 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:35:24 ; Search time 7744.85 Seconds
(without alignments)
9654.718 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgccctacagcaca.....gcaataacattctaaatatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	97.8	4.8	1101	9	CNS00EVL	AL069706 Drosophil
c 2	97.2	4.7	1101	9	CNS0039G	AL063921 Drosophil
c 3	95.6	4.7	1542	9	AG386981	AG386981 Mus muscu
4	95	4.6	1101	9	CNS0039G	AL063921 Drosophil
5	94.6	4.6	1758	9	CL509408	CL509408 SAIL_811_
c 6	93.6	4.6	1896	9	CG753083	CG753083 P048-1-C0
c 7	93	4.5	1202	8	CC262481	CC262481 CH261-167
8	92.4	4.5	1101	9	CNS0021J	AL061936 Drosophil
9	90.8	4.4	1101	9	CNS00EVL	AL069706 Drosophil
10	90.4	4.4	1608	9	CL118721	CL118721 ISB1-72J8
11	88.8	4.3	1067	6	CD386564	CD386564 AGENCOURT
c 12	88.8	4.3	1268	9	AG347098	AG347098 Mus muscu
c 13	88.8	4.3	1392	9	CG757503	CG757503 P052-4-C0
14	88	4.3	1277	8	CC253231	CC253231 CH261-180
c 15	88	4.3	1696	9	AG346840	AG346840 Mus muscu
16	87.6	4.3	1275	9	CL033318	CL033318 CH216-36F
c 17	87.2	4.2	887	9	AG526041	AG526041 Mus muscu
c 18	87	4.2	1275	9	CL033318	CL033318 CH216-36F
c 19	87	4.2	2087	9	AG333887	AG333887 Mus muscu
c 20	86.8	4.2	1348	9	CG749499	CG749499 P043-4-A0
c 21	85.6	4.2	1350	9	CL019486	CL019486 CH216-5G1
c 22	85.4	4.2	1780	9	AG320553	AG320553 Mus muscu
23	85.2	4.2	1320	9	CL103881	CL103881 ISB1-42C8
c 24	85	4.1	1539	9	AG340947	AG340947 Mus muscu
c 25	84.8	4.1	1254	9	AG349719	AG349719 Mus muscu
26	84.6	4.1	987	9	CNS014PQ	AL104456 Drosophil
c 27	84.6	4.1	1364	9	CG757970	CG757970 P053-2-C0
c 28	84	4.1	1101	9	CNS017KE	AL108152 Drosophil
29	83.8	4.1	822	7	CK416977	CK416977 AUF_IpInt
c 30	83.6	4.1	1489	9	AG350139	AG350139 Mus muscu
31	83.4	4.1	1101	9	CNS00FMC	AL070972 Drosophil
c 32	83	4.0	1211	9	AG349657	AG349657 Mus muscu
33	83	4.0	1592	9	CG750135	CG750135 P044-3-D0
c 34	82.8	4.0	1210	9	CG749728	CG749728 P044-1-C0
c 35	82.6	4.0	1506	9	AG278469	AG278469 Mus muscu
36	81.6	4.0	1092	9	CNS020K7	AL175696 Tetraodon
37	81.6	4.0	1101	9	CNS00EO7	AL069440 Drosophil
c 38	81.6	4.0	1355	9	AG346348	AG346348 Mus muscu
39	81.6	4.0	1745	9	AG338221	AG338221 Mus muscu
c 40	81.4	4.0	812	8	BH178455	BH178455 011_J_02-
c 41	81.4	4.0	812	9	CNS07KRM	AL615412 T3 end of
c 42	81.4	4.0	1050	9	CNS013NS	AL103090 Drosophil
c 43	81.4	4.0	1391	9	CG754863	CG754863 P050-2-G0
c 44	81.2	4.0	1094	9	CNS012FZ	AL101513 Drosophil
c 45	81.2	4.0	1715	9	AG288305	AG288305 Mus muscu

ALIGNMENTS

RESULT 1
CNS00EVL/c
LOCUS CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC: BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"
ORIGIN

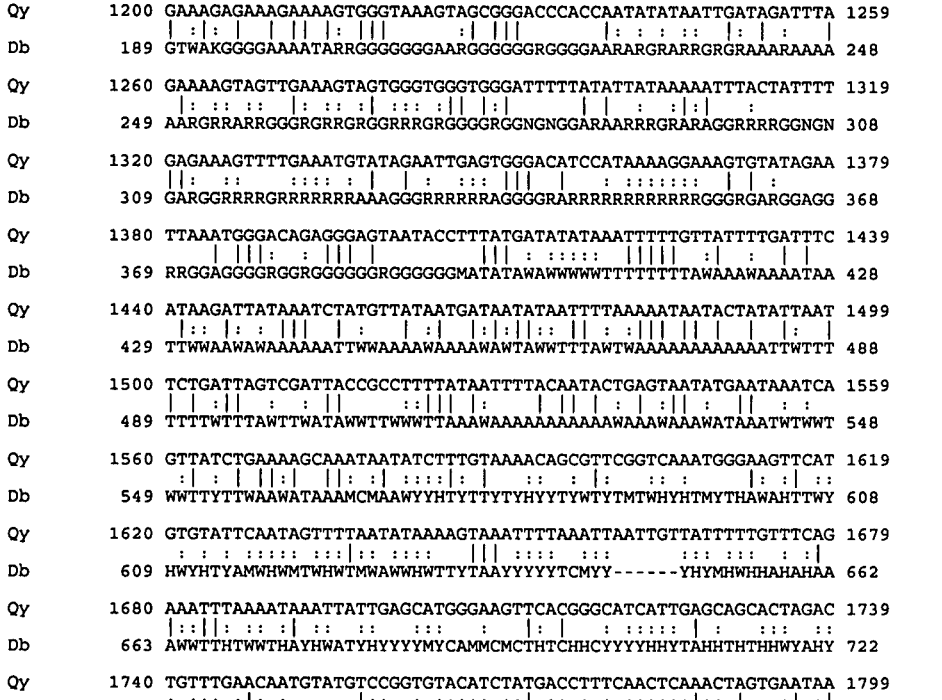
Query Match 4.8%; Score 97.8; DB 9; Length 1101;
Best Local Similarity 34.8%; Pred. No. 8.4e-08;

```
  Matches  195;  Conservative  102;  Mismatches  264;  Indels    0;  Gaps    0;

Qy       1138 TGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAA 1197
Db       1075 TAWATATATAAWTAATTWWTWTATATWTAWWWTATWATATATTAWTTTWTAAAATATAWW 1016

Qy       1198 AAGAAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATT 1257
Db       1015 TTAATAAATATATWTAWTAWWATATATAWATAWAAWTATAWAWTWWAATWAATWWATAWA 956

Qy       1258 TAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATT 1317
Db        955 WTTTAATAAWAATATWAWWWTWTTATTTTTTTTTTTTTWTTTWTATWTWTTTATATWWTW 896

Qy       1318 TTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAG 1377
Db        895 WAAAAWAWWAAWTWTATTTTTTWTTATWATTWATATTAWATTTTTTWTWTATTWWTATAT 836

Qy       1378 AATTAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATT 1437
Db        835 WATWWTATWTWTWTTTWATWTATWTATTTATWTWTWTTWTATTWTATWTATAWWWTWWAT 776

Qy       1438 TCATAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTA 1497
Db        775 WWWTAWAAWTATATATWWATATTATWWTWATTWTTWTTTTATTWTTTTAAWTTATWTATW 716

Qy       1498 ATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAAT 1557
Db        715 ATTWTTWATWTTTTTTTTTTTTTTTWTATTTWAAATTATTTAATTAWWWAATWAWTTATT 656

Qy       1558 CAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTC 1617
Db        655 TAATAAWTAATTAAATTATTAAAAATTTTATWAAAAAATATTTTTTTATATTAATWAWAT 596

Qy       1618 ATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTC 1677
Db        595 TTATAWATTAATATAWWTTWTAAATTWTTAATWTATTWWTTTAAATTATWTTTTTTTTTA 536

Qy       1678 AGAAATTTAAAATAAATTATT 1698
Db        535 WTAWAWWTWAAAWAATWAWWW 515


RESULT 2
CNS0039G/c
LOCUS       CNS0039G                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08K10"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             4.7%;  Score 97.2;  DB 9;  Length 1101;
  Best Local Similarity   18.8%;  Pred. No. 1.1e-07;
  Matches  109;  Conservative  261;  Mismatches  210;  Indels    0;  Gaps    0;

Qy       1133 TGAAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGA 1192
Db       1059 DRWAWDDGAGTWWTATWWWWWWWWWATWDTWWDKWWWWATAAKTDTAWTWWRTAWRADWAG 1000

Qy       1193 GTAAAAAGAAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGAT 1252
Db        999 RDRGAGKRDRDAATDADGAGRRDGGRKRKDKKDRKDGDDDKKGGKKKKAAKAAKWATKWW 940

Qy       1253 AGATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTA 1312
Db        939 DDWDWDKDWKWDGAKDRKADDDDGAGDKDDDGKGKDADDDTDGTKDDDDKDKWDDWDKAK 880

Qy       1313 CTATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTG 1372
Db        879 GTWGDATWAWAATDWWWWGWADADWWTWDAAADDWWADDRWDAWAWKWDDAWAWGARTAD 820

Qy       1373 TATAGAATTAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTT 1432
Db        819 RRDWGDRAGKRGGARKRRDRKRADDKRDAADDRDDAATWTTWTTTTTRDTDDWKWKTDTW 760

Qy       1433 TGATTTCATAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTA 1492
Db        759 TRWAADRTWDRDDDDDRDRAGTAGRKWRRTWKRRWKRRDTRWDDADADDTARDDRRRRGD 700

Qy       1493 TATTAATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAA 1552
Db        699 DGADAGKGKKTGRKRRRRDRATWDRTDAWWADAAWWTTTDTDTDDWDKRDRRRKGARRRR 640

Qy       1553 TAAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGGGA 1612
Db        639 RTTARAAWDWWTWKAWDWAKWDWKTRADRWDRWAADTWTDARKADRDWAKARAWRARRDR 580

Qy       1613 AGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTT 1672
Db        579 ARAARADRRWTTKGKTTTATWTTWAARAAWWAWWAWATTTATWTTTWTTTWTTTTTTTTT 520

Qy       1673 GTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTT 1712
Db        519 TTWTTTAAWWWAAWWTATWAAWTAAAWAAAAAAAAWAATTT 480


RESULT 3
AG386981/c
LOCUS       AG386981                1542 bp    DNA     linear   GSS 03-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-201G10.TJ, genomic survey
            sequence.
ACCESSION   AG386981
VERSION     AG386981.1  GI:47998186
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
```

```
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1542)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1542
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-201G10.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             4.7%;  Score 95.6;  DB 9;  Length 1542;
  Best Local Similarity   44.3%;  Pred. No. 2.1e-07;
  Matches  420;  Conservative    0;  Mismatches  523;  Indels    6;  Gaps    1;

Qy        750 ATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAAGTTACTAATAA 809
              |||| | | |   ||    ||| | || |||| || |  || ||| |||  || | |  |
Db       1103 ATAAGAAAATTATATAATAAATAATATAAAAAATATAATTAATAAATAAANTATTTAATA 1044

Qy        810 ATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAGAGCTTGCTGCTGTGTGTTTAG 869
              | |  | |    |  ||| | | |  | |||||| || |    |       |     ||
Db       1043 AAATAATAANAAATAAAAAAAATAAAATATAAAAAAAATAATATTTAAAAATAAANATAA 984

Qy        870 TTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGATCTAAAGCACATAGAAATTTA 929
              |             || |  |   ||  |||   ||   ||  ||   | |   | | |||
Db        983 TATAATAATATATATATAATATTAATAAATAATAATATATAAAATAAANAATTATAATTA 924

Qy        930 GTACAGGTTAAAACTTTTACAAGAATTTATATTAAACGAAAATCATTTTATAACATGTCT 989
              || |   ||||| | ||| || ||     | | |   | ||  || |     | || | |
Db        923 ATAAATAATAAAATTATTAAAAAAAAAATAAATTATAAATAAATATATATATAAATATAT 864

Qy        990 CTCGGCTGTCATTATAATAGGGATCACTTACTGATCATCCATTAA------AACCTTGTT 1043
              |    | | |||| ||||   |  | | |      |   |||||      ||   |  |
Db        863 TTAAATTATAATTAAAATAATAAATAATAATATTATAAAAATTAATTAAATAAATATAAT 804

Qy       1044 AAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGGACAATGTCTCTTTGAA 1103
              |||| ||| | ||  | |||||| |   || ||  |  | ||    |   | ||
Db        803 AAAAAAAAATAAAAAAAATAAAAAAAAATAAAAATATNAAAATAAAAAAAAATAAATTAA 744

Qy       1104 AAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATATGGATTGGACACGGAG 1163
              |||  ||  |  || |              |||  ||   | |    |  | |   |
Db        743 TAAAATAAAATATAATAATAAAAAAATATAAAAAAATAATAAAAATAAAAATAAAAAAAT 684

Qy       1164 ACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGAAAGAAAAGTGGGTAAA 1223
              | ||| ||| || | || |   |   | | |||| |  |    ||| |||| |   ||||
Db        683 AATAATAAATATATTTATAAATAAAAATAATAAATAATATTTTAAATAAAAATAAATAAA 624

Qy       1224 GTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGT 1283
              ||          |  |   | |||       |  |  ||   |     | | |  | |   |
Db        623 ATAAAAAATAATAAAATATTTAAATATTATTAAATNTAATTATTAAAAATAAAAATATAT 564

Qy       1284 GGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAA 1343
                     |  |||| |||||||| |    |   | | ||| | |  | | ||    |
Db        563 TATATATAAAAATATAATATAAAAAATATAAAAAATAATAAAATATAAATTTATAATATA 504

Qy       1344 TTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAATA 1403
              | |        |  | |      |   |      | ||||    | | |      ||| |
Db        503 TAAAAAATATAAAAAAAATATTATTTATATTTATAATATAAAAAAAAAATATTTTATAAAA 444

Qy       1404 CCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGATTATAAATCTATGTTATA 1463
                | ||    | | |||||| |   || |       | || |||   | |  |   ||||
Db        443 ATTATAAATAAAAAAAATTTATAAAATATATTAAAAAAAAAATTTATATTAAAATATATA 384

Qy       1464 ATGATAATATAATTTTAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGCCTTT 1523
              || | || | || | |||||    || || || | ||   | | |      |        |
Db        383 ATAAAAAAAAAAATATAAAATATTTATTAAATAATTTAATAAAATTAANAAAATTAAAAT 324

Qy       1524 TATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCTGAAAAGCAAATAATATC 1583
                 ||  |   || |   | || | | ||   |  |  |||     ||  | || | ||
Db        323 ATATATATATAAAAATAAATTATTTTAAAATTAAAATATATAATTTAAAAATATTAAATA 264

Qy       1584 TTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAATATAA 1643
              | |  || |       | | ||   || |   |   ||||  |||    | | | | |
Db        263 AATATTTAAAAAAAAAATATAATATTAAAAATAATTTAATATTAAATTAAATAATTTTTA 204

Qy       1644 AAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTAAAATAA 1692
              ||   | | ||| | ||   || ||     || |  || ||  |  ||
Db        203 AAAATATATATAATTAAAAATTTTTTAAATATTAANTAAATTNTATAAA 155

RESULT 4
CNS0039G
LOCUS       CNS0039G                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08K10"
```

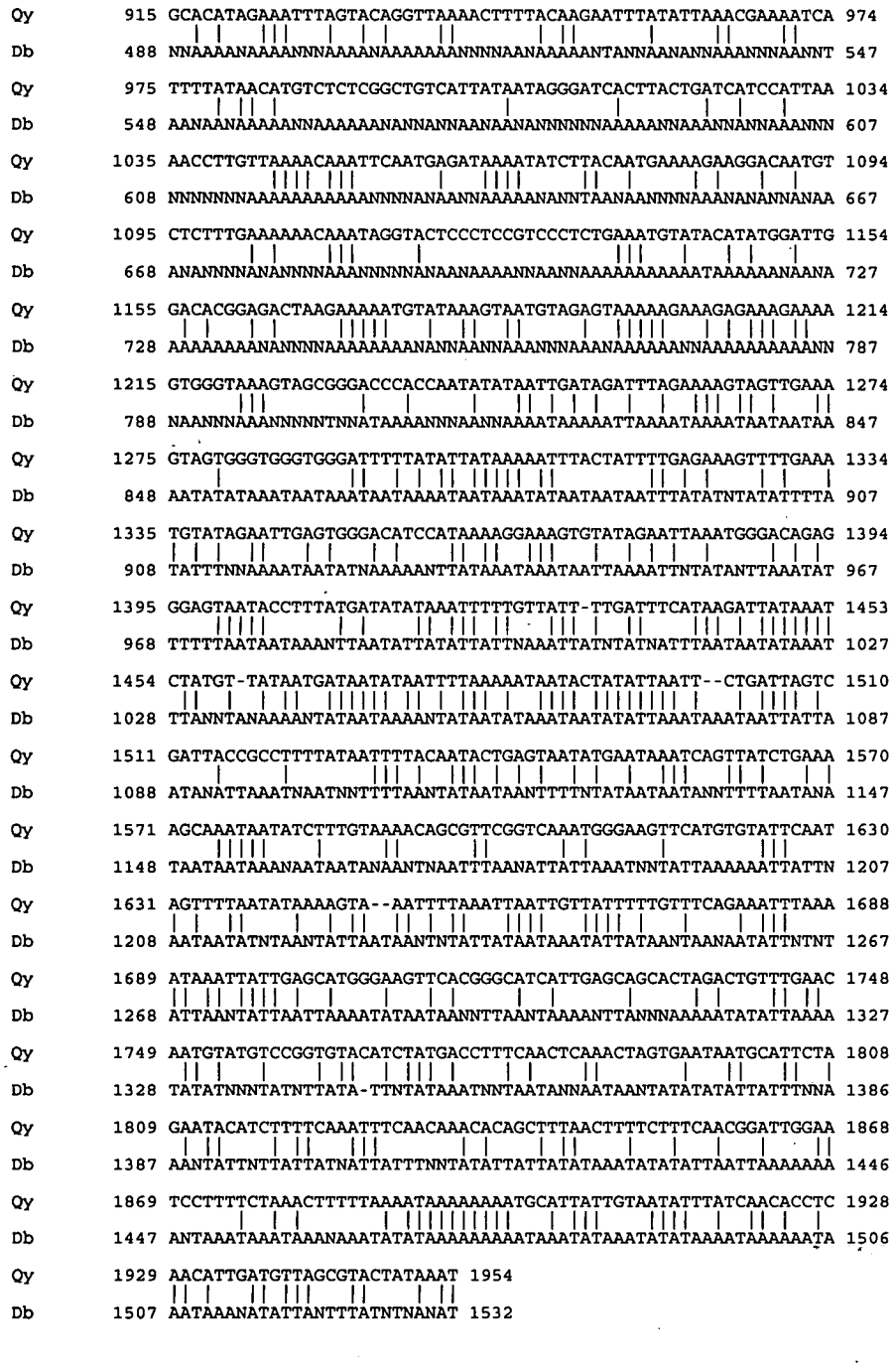
/clone_lib="RPCI-98"
/note="end : TET3"
ORIGIN

Query Match 4.6%; Score 95; DB 9; Length 1101;
Best Local Similarity 20.2%; Pred. No. 2.7e-07;
Matches 173; Conservative 325; Mismatches 348; Indels 9; Gaps 2;



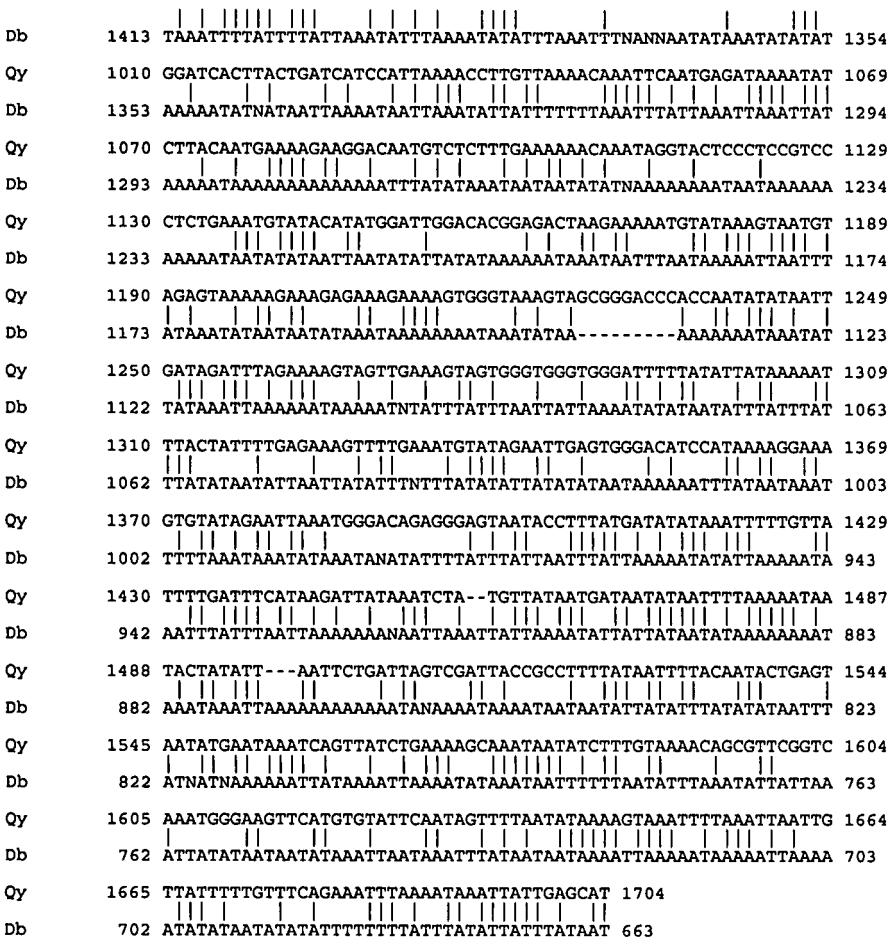
RESULT 5
CL509408
LOCUS CL509408 1758 bp DNA linear GSS 01-APR-2004
DEFINITION SAIL_811_H11.v3 SAIL Collection Arabidopsis thaliana genomic clone
SAIL_811_H11.v3, genomic survey sequence.
ACCESSION CL509408
VERSION CL509408.1 GI:46006728
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1758)
AUTHORS Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B., Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
TITLE A high-throughput Arabidopsis reverse genetics system
JOURNAL Plant Cell 14 (12), 2985-2994 (2002)
MEDLINE 22356987
PUBMED 12468722
COMMENT Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS836276; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not single contiguous sequences.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .1758
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_811_H11.v3"
/clone_lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"
ORIGIN

Query Match 4.6%; Score 94.6; DB 9; Length 1758;
Best Local Similarity 34.2%; Pred. No. 3.2e-07;
Matches 502; Conservative 0; Mismatches 957; Indels 7; Gaps 5;



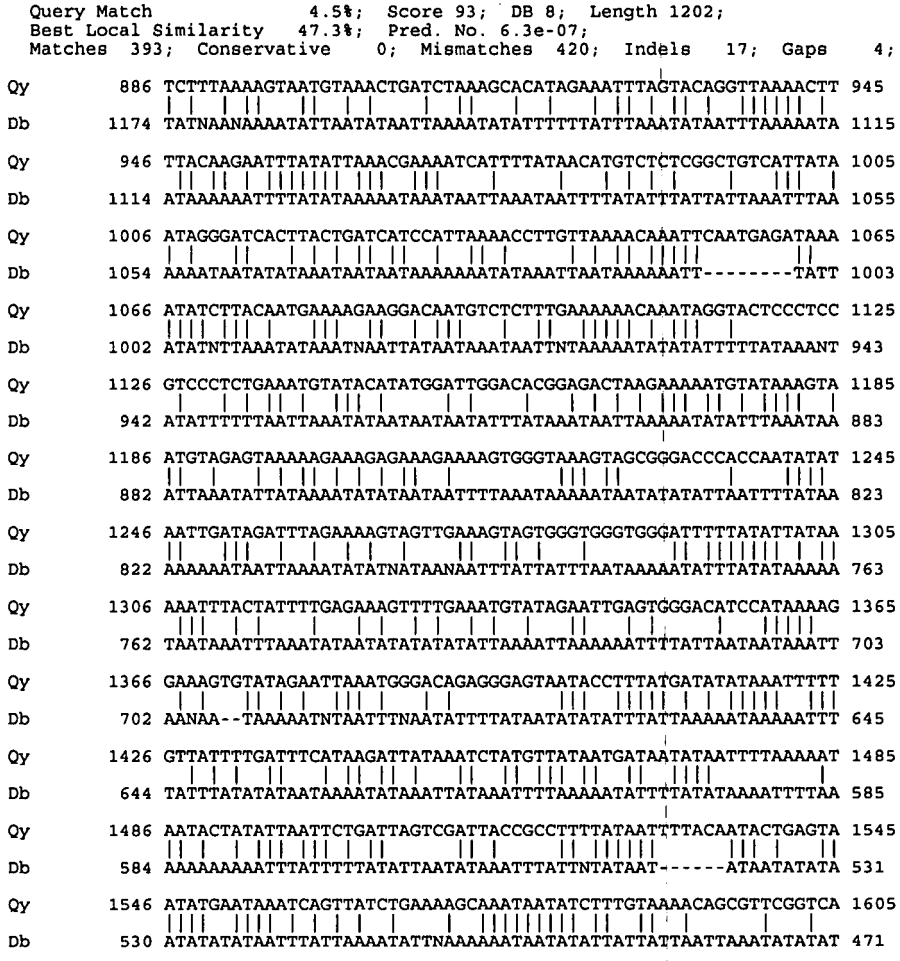
RESULT 6
CG753083/c
LOCUS CG753083 1896 bp DNA linear GSS 24-OCT-2003
DEFINITION P048-1-C01.za Ppa EcoRI BAC Library Pristionchus pacificus genomic, genomic survey sequence.
ACCESSION CG753083
VERSION CG753083.1 GI:37977199
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1896)
AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K., Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .1896
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."
ORIGIN

Query Match 4.6%; Score 93.6; DB 9; Length 1896;
Best Local Similarity 44.1%; Pred. No. 4.9e-07;
Matches 520; Conservative 0; Mismatches 645; Indels 15; Gaps 4;



RESULT 7
CC262481/c
LOCUS CC262481 1202 bp DNA linear GSS 13-MAY-2003
DEFINITION CH261-167M9_Sp6.1 CH261 Gallus gallus genomic clone CH261-167M9, genomic survey sequence.
ACCESSION CC262481
VERSION CC262481.1 GI:30607397
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1202)
AUTHORS Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 30
High quality sequence stop: 105.
FEATURES Location/Qualifiers
source 1. .1202
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-167M9"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI; CH261 Female Chicken library - for library and clone ordering information: http://www.chori.org/bacpac"
ORIGIN

Query Match 4.5%; Score 93; DB 8; Length 1202;
Best Local Similarity 47.3%; Pred. No. 6.3e-07;
Matches 393; Conservative 0; Mismatches 420; Indels 17; Gaps 4;



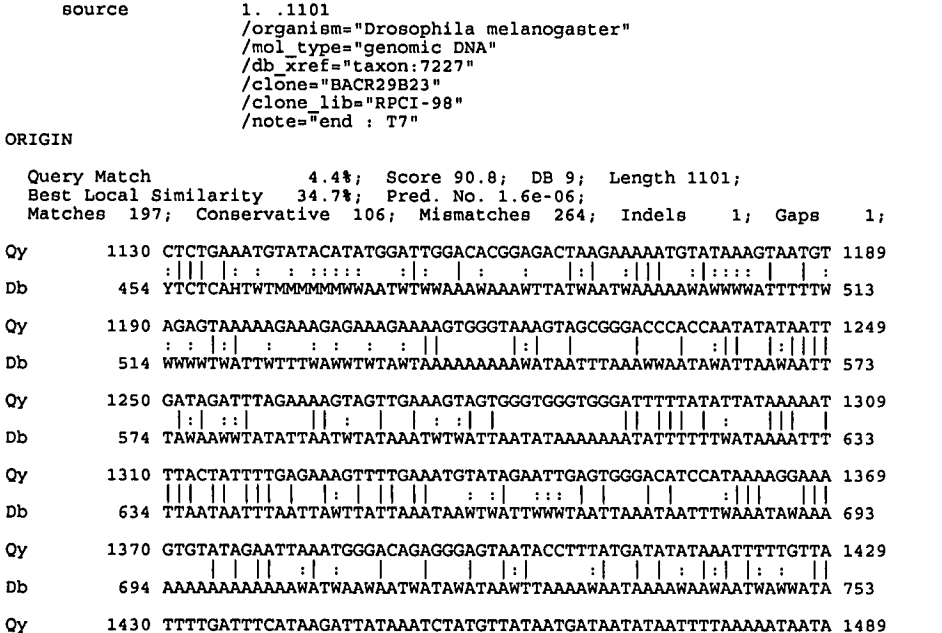
```
Qy       1606 AATGGGAAGTTCATGTGTATTCAATAGT-TTTAATATAAAAGTAAATTTTAAATTAATTG 1664
Db        470 AAAAATAATAATATATATTTTAAATTTTATTTAATTTTAAATAAAANTNAAAAATAATTT 411

Qy       1665 TTATTTTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGGGAAGTTCA 1714
Db        410 TTTTATATAAAATATTAATATATATATAATAAATAATTATTATAAATTAA 361


RESULT 8
CNS0021J
LOCUS       CNS0021J                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL061936
VERSION     AL061936.1  GI:4940214
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR05N11"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             4.5%;  Score 92.4;  DB 9;  Length 1101;
  Best Local Similarity   39.4%;  Pred. No. 8.2e-07;
  Matches  349;  Conservative   65;  Mismatches  469;  Indels    2;  Gaps    2;

Qy        868 AGTTGTTGTGAGCTCATTTCTTTAAAAGTAATGTAAACTGATCTAAAGCACATAGAAATT 927
Db        138 AGTGTTTTCATGAAAGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANTCNNNNCNNAAN 197

Qy        928 TAGTACAGGTTAAAACTTTTACAAGAATTTATATTAAACGAAAATCATTTTATAACATGT 987
Db        198 NAAAAANAAAAAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 257

Qy        988 CTCTCGGCTGTCATTATAATAGGGATCACTTACTGATCATCCATTAAAACCTTGTTAAAA 1047
Db        258 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAMAAAAAAA 317

Qy       1048 CAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGGACAATGTCTCTTTGAAAAAA 1107
Db        318 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 377

Qy       1108 CAAATAGGTACTCCCTCCGTCCCTCTGAAATGTATACATATGGATTGGACACGGAGACTA 1167
Db        378 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 437

Qy       1168 AGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGAAAGAAAAGTGGGTAAAGTAG 1227
Db        438 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAANAAAAAAAAAAAAAAAAAAAAAAAA 497

Qy       1228 CGGGACCCACCAATATATAATTGATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGT 1287
Db        498 AAAAAAAANAANAAAAATATAATTTAWTTTTTTWTTAATTAWTTTTTTTTTTTTTTTTT 557

Qy       1288 GGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAATTGA 1347
Db        558 TWTTAATTTTAATTTTTAAWAWAAATTTAATAAAAWAWTWTTTAWTTTTAATWTAAWWAA 617

Qy       1348 GTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAATACCTT 1407
Db        618 AAAAAAWTTTTAAAWAWTWTTTTTTTTWTTAATTAATAWAAAAAAAWATAWAAWTTWTTTA 677

Qy       1408 TATGATATAT-AAATTTTTGTTATTTTGATTTCATAAGATTATAAATCTATGTTATAATG 1466
Db        678 TWWATWAAATWAAATTTTAAAAAWTTTTTTAWTTTTTTTWTTTAAAAWTTTATTAWTTTT 737

Qy       1467 ATAATATAATTTTAAAAATAATACTATATTAATTCTGATTAGTCGATTACCGCCTTTTAT 1526
Db        738 AAAAWATWAWTWWTAAATTTTTTAWATAWTAAWATTAWAAAATTTTTTAATTTAATTTTT 797

Qy       1527 AATTTTACAATACTGAGTAATATGAATAAATCAGTTATCTGAAAAGCAAATAATATCTTT 1586
Db        798 TTTAATTAAAAAAAAAAAAAAAAAAAAWAAATAAAAAAATAAAAAAAAAAAWAAAWWTTTTT 857

Qy       1587 GTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAG 1646
Db        858 TTAAATTATTTWTTATTTAAAAATTTTAATAATSATTTAAWTATGTTTTTTAAGTWTAA- 916

Qy       1647 TAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAATTTAAAATAAATTATTGAGCATGG 1706
Db .      917 AAAWTGKGAATGKKKKGKTARKKKGRKGTTAARAATDGTAAGAKRATAAKKTKTKAATKK 976

Qy       1707 GAAGTTCACGGGCATCATTGAGCAGCACTAGACTGTTTGAACAAT 1751
Db        977 TAAAAAADKAKGAAAAATAWGAARGKATGAAAAAGTGKDAATATT 1021


RESULT 9
CNS00EVL
LOCUS       CNS00EVL                1101 bp    DNA     linear   GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION     AL069706.1  GI:4949849
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
```

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

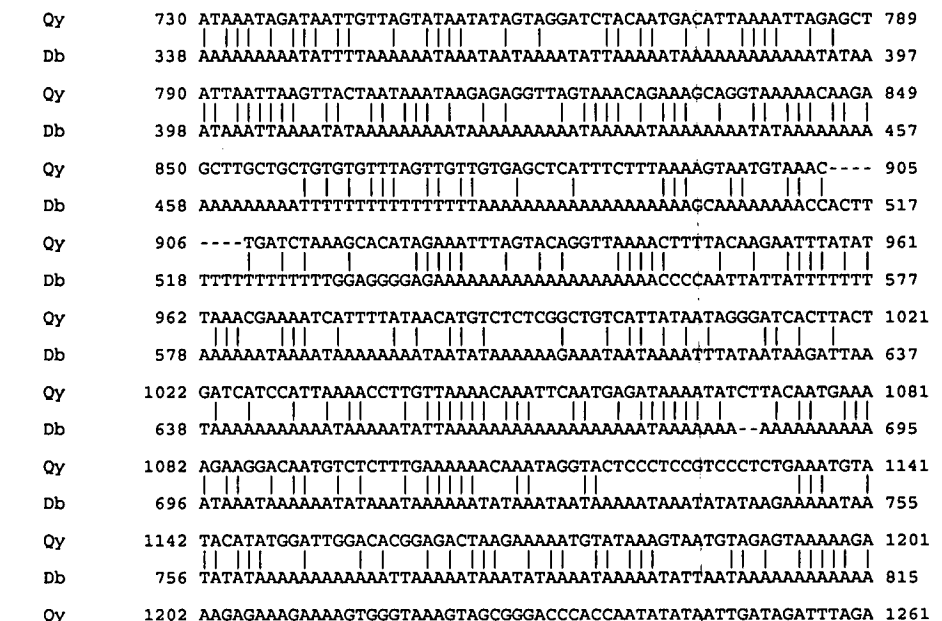
Query Match 4.4%; Score 90.8; DB 9; Length 1101;
Best Local Similarity 34.7%; Pred. No. 1.6e-06;
Matches 197; Conservative 106; Mismatches 264; Indels 1; Gaps 1;

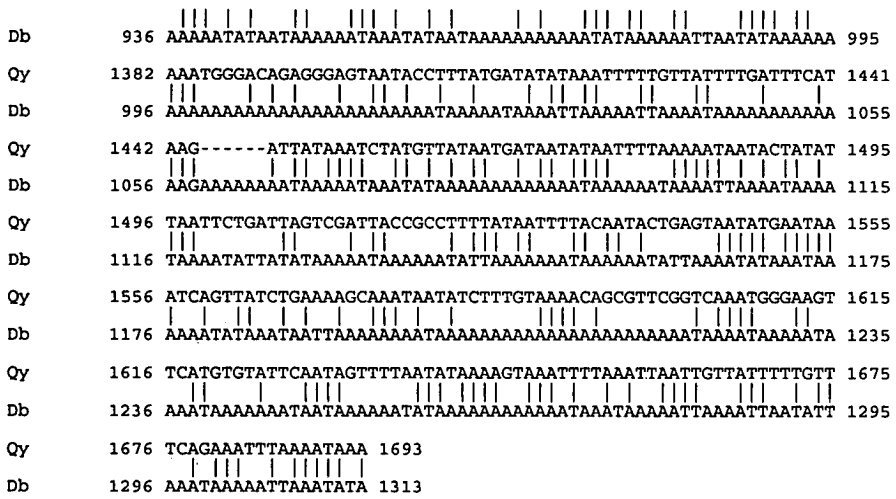


RESULT 10
CL118721
LOCUS CL118721 1608 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-72J8_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-72J8, genomic survey sequence.
ACCESSION CL118721
VERSION CL118721.1 GI:40612356
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1608)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 252
High quality sequence stop: 345.
FEATURES Location/Qualifiers
source 1. .1608
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-72J8"
/clone_lib="ISB1"
/note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"

ORIGIN

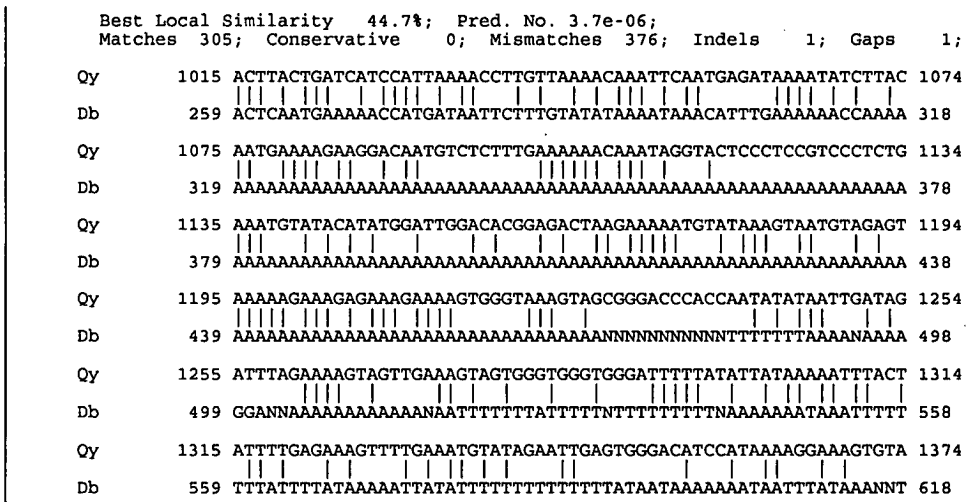
Query Match 4.4%; Score 90.4; DB 9; Length 1608;
Best Local Similarity 45.6%; Pred. No. 1.9e-06;
Matches 446; Conservative 0; Mismatches 516; Indels 16; Gaps 3;





RESULT 11
CD386564
LOCUS CD386564 1067 bp mRNA linear EST 30-MAY-2003
DEFINITION AGENCOURT_14286445 NIH_MGC_173 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CD386564
VERSION CD386564.1 GI:31222277
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1067)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDKM74 row: e column: 02
High quality sequence start: 8
High quality sequence stop: 462.
FEATURES Location/Qualifiers
source 1. .1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic trophoblasts, made from WA01 stem cells"
/lab_host="DH10B TonA"
/clone_lib="NIH_MGC_173"
/note="Vector: pDONR201; Site_1: attP2; Site_2: attP1; LIBR_PRIMING - oligo dT; METHOD - full-length enriched; LIBR PROVIDER - Bradfield"
ORIGIN

Query Match 4.3%; Score 88.8; DB 6; Length 1067;



RESULT 12
AG347098/c
LOCUS AG347098 1268 bp DNA linear GSS 02-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-142I02.T7, genomic survey sequence.
ACCESSION AG347098
VERSION AG347098.1 GI:47920408
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1268)
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

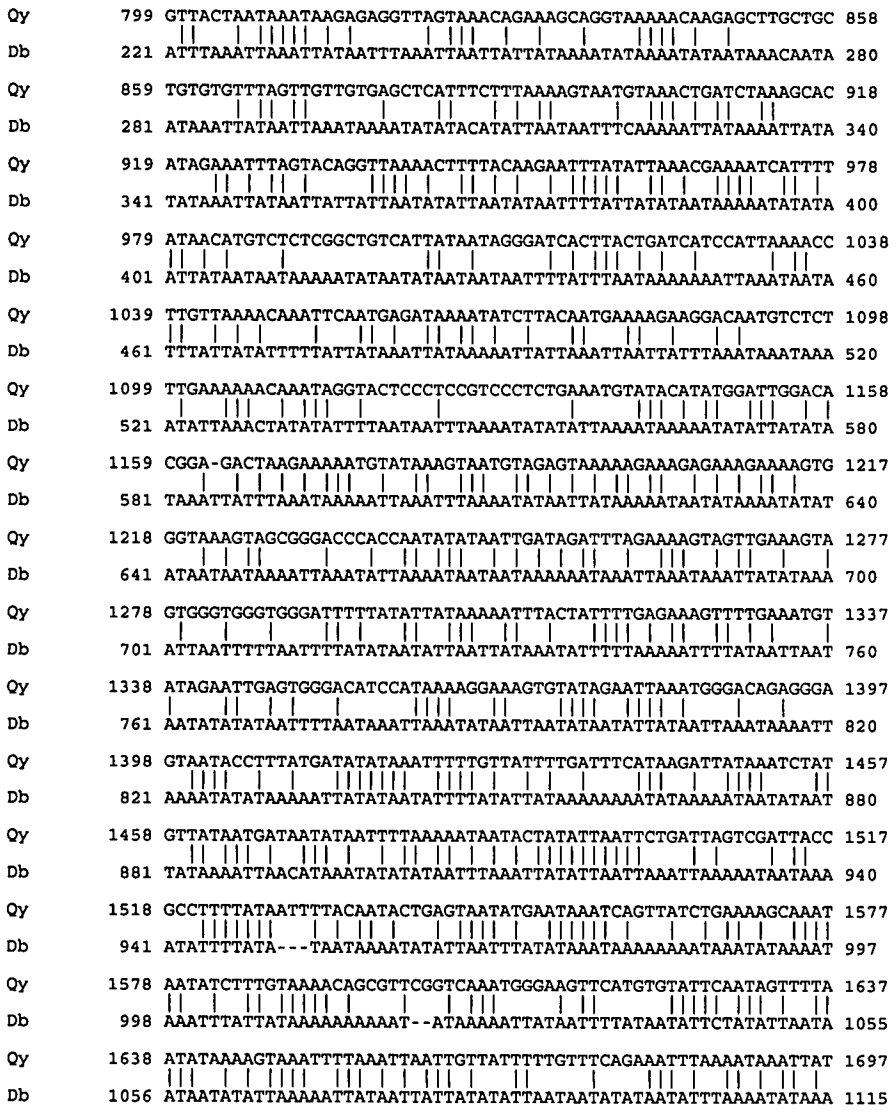
```
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1268
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-142I02.T7"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             4.3%;  Score 88.8;  DB 9;  Length 1268;
  Best Local Similarity   45.5%;  Pred. No. 3.7e-06;
  Matches  471;  Conservative    0;  Mismatches  551;  Indels   14;  Gaps    4;

Qy       1030 ATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATATCTTACAATGAAAAGAAGGAC 1089
              ||||||   |  | | || ||||  |   | ||| ||||  | |  ||  ||| ||  |
Db       1215 ATTAAATAATNATAATAAAAAATATATATAAATATAATAATTAAATATATAAAAAAATAT 1156

Qy       1090 AATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAAT---GTATACAT 1146
              |||   | | |  |||   |||    ||                |  ||||    ||  |
Db       1155 AATAATTATATATAAATTAAAAATAATAAATTAATAAATTATAAAAAATAAAATAATTAA 1096

Qy       1147 ATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAG 1206
              ||  | |    |   | |   |  |  |  ||||||  ||   | | || ||| ||  |
Db       1095 ATTAAATTATTATTAATAAATAATATTAAATATAAATAAAAAAAAATTATAAATAATTAT 1036

Qy       1207 AAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAGATTTAGAAAAGT 1266
              ||  |||  |   |||  ||      |   |   ||| ||||| | |  |||| |||| |
Db       1035 AATAAAATATAATTAATATAAAAATAATAAAATATAAATAATATAAAATTTTATAAAATT 976

Qy       1267 AGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACTATTTTGAGAAAG 1326
              | ||  ||  |     |   |   | |  | | ||| |||   | |||          |
Db        975 ATTTATAAAAATAATATTTATAAAAAATAATTTTAATATAAAAAAATTATAAAATATATA 916

Qy       1327 TTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAATTAAATG 1386
               |   ||   | |  |     |     | |   |||||   | |    | | |||  ||
Db        915 ATAATTAAAAAAAAATAAATAAAAAAAAAAATAATAAATATATATATAAAATAATATAAAA 856

Qy       1387 GGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTGATTTCATAAGAT 1446
                |   |   |  |||   ||| | ||||||||||||     |  | |  |    | ||
Db        855 TTAATAATATAAAAATTATTTTTTTATATATAAATTAAAATATTATATAANAAAAAAATA 796

Qy       1447 TATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAATTCTGATT 1506
                |||  |||       || |||   | ||   ||| | | || |  || ||||   ||
Db        795 ATAAAAAATATTAATATATAATATATTTATAAAAAATAAATTATTTAATAAATTAAAATA 736

Qy       1507 AGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATGAATAAATCAGTTATCT 1566
              | |  | ||     ||   ||  | | || | | |   | || | ||||| |   ||
Db        735 AATAAAATAATTTATTAATAAAAATAATAAAAATAAAATAAATAATTAAATAAAAAATAA 676

Qy       1567 GAAAAGCAAAT---AATATCTTTGTAAAACAGCGTTCGGTCAAATGGGAAGTTCATGTGT 1623
               |||    |||   |  || | | || || |         | | |      | || |
Db        675 AAAATAATAATTAAATAATATATATATAATAAAAANTAATAATAAAAATAATTAAAAATA 616

Qy       1624 ATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTTTTGTTTCAGAAAT 1683
```

```
              ||   ||| | | || | | || |||| | |||||||||  | |          | ||||
Db        615 ATATTATAATATAAAAAAATAAATAAAATATAAATTAATAATAAAAAAATAAAAAAAAAT 556

Qy       1684 TTAAAATAAATTATTGAGCATGGGAAGTTCACGGGCATCATTGA---GCAGCACTAGACT 1740
               ||||||| ||      |  ||   ||    |    |  ||  |           || | |
Db        555 ATAAAATATATATAATAATATTATAAAAAAAAAAAATAATATAAAAATTATAATTTAAAAT 496

Qy       1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
               | | ||  |   || |     |  |       ||     ||| ||  |||  |   |||||
Db        495 ATATTAAATAAAAATATAATATTAAATATAATTTAATTTAAAAAAAAAAAAAAAAATAAG 436

Qy       1801 GCATTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTT-----T 1855
                 | || |     | |  |||   || ||| |     |||   |||| ||
Db        435 AGTAAAAATAAAATTAGGTATTTAATAAAAAAAAATATTTATTTTTTTTTTNTTAAAAAA 376

Qy       1856 CAACGGATTGGAATCCTTTTCTAAACTTTTTAAAATAAAAAAAATGCATTATTGTAATAT 1915
               ||   |||  ||   |||| |||| |  | ||||||||| |||   | |   || |||
Db        375 AAAAAAATTATAAATATTTTATAAAATAATAAAAATAAAAATAATATTTAAAATTATTAT 316

Qy       1916 TTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACT 1975
               | |  || |    || ||  || | |   ||    || ||  |    || |    ||
Db        315 TAAAAAAAAAAAAAATATATATATAATTTTAAATAAATTATTTTAAATTAATATAATAAA 256

Qy       1976 ATCATCACATCAATCTTACACCACAAACCTTGAGCTTAATTTTTCTACTTATTCTCAGCA 2035
                |    | ||  | |    | | |     |  | | ||  | |  | ||   | |
Db        255 TAATTATATTAAAATTAAATTAATATAAAAAAAAATAATTTAATAAATAAAATAATAATA 196

Qy       2036 ATAACATTCTAAATAT 2051
              |||| |||    ||||
Db        195 ATAAAATTAATTATAT 180
```

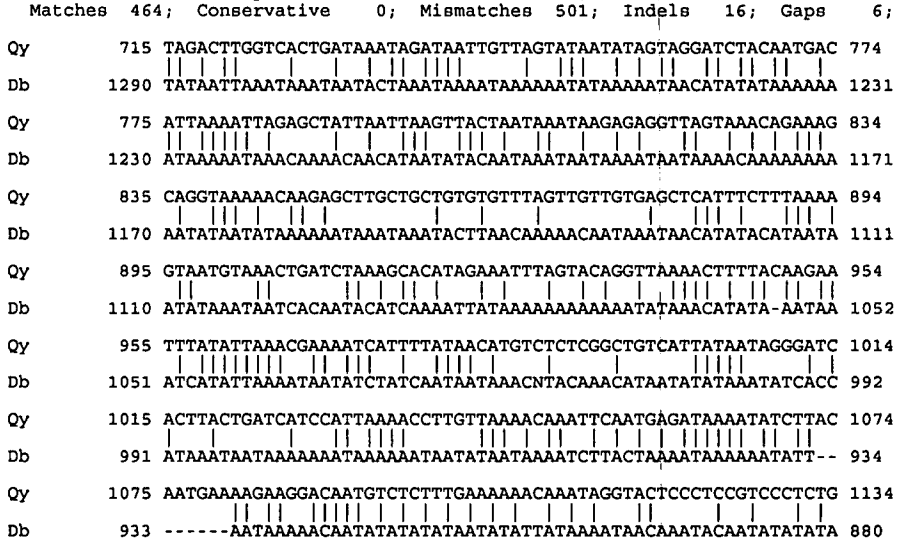
```
RESULT 13
CG757503/c
LOCUS       CG757503                1392 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P052-4-C08.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG757503
VERSION     CG757503.1  GI:37986131
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1392)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .1392
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             4.3%;  Score 88.8;  DB 9;  Length 1392;
```

```
  Best Local Similarity   43.8%;  Pred. No. 3.7e-06;
  Matches  543;  Conservative    0;  Mismatches  672;  Indels   24;  Gaps    3;

Qy        481 ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACACTCTATCCACGTACTATACTTT 540
              ||  | |   ||  ||| | | |||         | | |   |||  |   |  ||| |
Db       1359 ATAATAAAATATATAAAATAATAAATAAATTAAAATAAAAATATTTAAAAATAATAATAA 1300

Qy        541 TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAATAATGCCCATCCAAGGATAAGT 600
              |    ||   |   || ||   | | |   | | | ||    ||  ||  | ||
Db       1299 TAAAAAAAAAATAANTAATATAAATAAAANAATAATATTTATAATAATAAAAAAAAAAAA 1240

Qy        601 AAAATTCCCGTTTAACCAGTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGT 660
              |  |||       |||  |       || | | |     | | |   || |  |   |  |
Db       1239 ATTATTAAAAAATAAATATAAAAAAAAAAAAAAAAAAAAAAAAAATATAAAAATAAATATAAT 1180

Qy        661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
              || |  |  |    |  |     |||    ||| ||  |    |  | | |     |
Db       1179 AAAAAATAAAATAAAATATTATATTAAAATAAATATAAAAAATNAAATATAAAAAAATAA 1120

Qy        721 TGGTCACTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAA 780
              |    ||||| | | ||    || ||  | | |  |  |     ||| | |  | |
Db       1119 AAAATAAAAATAAAAATAAAAAAAATAATAATAAAAAAAATAAAAATTAATAATAAAATA 1060

Qy        781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
              |  | |   | ||| || ||| ||||| |||| | |  | |  ||  |  ||  |   |
Db       1059 AAAAAAATATTAAATAAAATTATTAATATATAAAAAAAATAAAAAATAAATAAATAATAA 1000

Qy        841 AAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAAGTAATG 900
              ||  | ||      |              | |    |    |         ||||  |||
Db        999 AATAAAAATAAATTAAAAATATAAAAAAATATATAAAAAAAAAAAAAAAAAAAAAAATT 940

Qy        901 TAAACTGATCTAAAGCACATAGA-----------AATTTAGTACAGGTTAAAACTTTTAC 949
              | |   || ||||  | | | |           ||  |  || |  |||||  |   |
Db        939 AATATATATATAAAATATAAAAANAAAAATAAATAAAATTATAAAAATTAAATATAAAAT 880

Qy        950 AAGAATTTATATTAAACGAAAATCATTTTATAACATGTCTCTCGGCTGTCATTATAATAG 1009
              ||  || |||||||| |  ||||  |    | ||  |          | | | |  ||||
Db        879 AATAAATTATATTATAAAAAAAAAAAAAAAAAAAAATTAAAAAATATATATTAATTAAATAA 820

Qy       1010 GGATCACTTACTGATCATCCATTAAAACCTTGTTAAAACAAATTCAATGAGATAAAATAT 1069
                |  |    |      |   |  ||||   |   |||| | ||  ||  | |||  |||
Db        819 ATAAAATAAATAATAAAAATAAAAAAAAAATAAAAAAATATATATAAATAAATATAATTA 760

Qy       1070 CTTACAATGAAAAGAAGGACAATGTCTCTTTGAAAAAACAAATAGGTACTCCCTCCGTCC 1129
              |||  |   ||| ||  | ||    |   | |||||| |||||   |
Db        759 ATTATTAAATAAATAAAAAAAAAAAAATAAATAAAAAAAAAAATAAAAATATAAAAAAAAA 700

Qy       1130 CTCTGAAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGT 1189
                |||   | |    |  ||   | |   | |   | |   ||  ||||||  ||
Db        699 AAAATAAAATAAAAATATTAAATAAAATATAAATAAATAAATTAAAATATAAAAAAAATA 640

Qy       1190 AGAGTAAAAAGAAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATT 1249
              | | ||||||  || | |   | ||       |||  |          |  || ||| ||
Db        639 AAAATAAAAAATAAAATATTTATAAAAATAAAAAAAAATNTATAATAAAAAAATAAAAAA 580

Qy       1250 GATAGATTTAGAAAAGTAGTTGAAA----GTAGTGGGTGGGTGGGATTTTTATATTATAA 1305
              | | || || |||  |       ||     || |   |   |    ||| |  ||||| |
Db        579 AAAAAATATAAAAATTTTTAAATAAATTTATATTAATTTTATTTATTTTATTAATTATTA 520

Qy       1306 AAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAG 1365
              |   |   |||||  | ||  | |  |||   | | ||  |     | |   | |||
Db        519 ATTATATATATTTATATAATATATAAAAAAAAAAAAAAAAAAAATAATTAAAAAAAAAAATA 460

Qy       1366 GAAAGTGTATAGAATTAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTT 1425
              ||     | | |    |||   | | |   ||||   | |    | |  |||  |
Db        459 TAATTAAAAAAAATAATAATAAAAAAAAAAAATAAATAAAATAAATTAAAAATAATAATA 400

Qy       1426 GTTATTTTGATTTCATAAGATTATAAATCTATGTTATAATGATAATATAATTTTAAAAAT 1485
                ||  |||| | |   |  |  ||   | | | ||| |||||||||   | |||||
Db        399 AAATATTATATTTAAAATAAAAAATAAATAAAATAAAAATTATAATATAAAATAAAAAAA 340

Qy       1486 AATACTATATTAATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTA 1545
              ||   |||| | | | |   |     ||||      | || | || || | ||   |
```

```
Db        339 AAATATATAATTAATTTATATTTATTATTATAAATATATAAATTTATAAATTATATATAT 280

Qy       1546 ATATGAATAAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCA 1605
                |  | | |   |  ||| | ||||  |   ||||  |||  || | |         |
Db        279 TAAAAATTTATATAAATATATAAAAAATATTAAATAAATTTAAAATAAAAAAAAAAAAAA 220

Qy       1606 AATGGGAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGT 1665
              ||    |  || || | |          |  ||| ||  ||  ||| |||||  ||||
Db        219 AAATATATATTAATTTTT---------TAATAAAATTTAAAAAAAATTTAAAATAATATA 169

Qy       1666 TATTTTTGTTTCAGAAATTTAAAATAAATTATTGAGCAT 1704
              |||| || |   | |||||  |||| |||  |   |  ||
Db        168 TATTATTATAAAAAAAATTAAAAAAAAAAANAAAAAAAAT 130


RESULT 14
CC253231
LOCUS       CC253231               1277 bp    DNA     linear   GSS 13-MAY-2003
DEFINITION  CH261-180N11_RM1.1 CH261 Gallus gallus genomic clone CH261-180N11,
            genomic survey sequence.
ACCESSION   CC253231
VERSION     CC253231.1  GI:30589981
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1277)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 24
            High quality sequence stop: 85.
FEATURES             Location/Qualifiers
     source          1. .1277
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-180N11"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN

  Query Match             4.3%;  Score 88;  DB 8;  Length 1277;
  Best Local Similarity   43.9%;  Pred. No. 5.2e-06;
  Matches  475;  Conservative    0;  Mismatches  600;  Indels    6;  Gaps    3;

Qy        619 GTTTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGACAA 678
              || |||   | | | ||| |     |  | | ||   |  |                ||
Db         41 GTGTGTGTGTGTGTGTGTGTGTGTGTATGGGGGGGTGAGGGGGGNNNNNNNNNNNNAA 100

Qy        679 GAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGA 738
              ||  |||     ||       |      || | |   || |       | | |  |||| |
Db        101 GAACCTTTAAAAAANNCNCAAAAAAAATACATATTTTATAAAATAATATTAAATAATATA 160

Qy        739 TAATTGTTAGTATAATATAGTAGGATCTACAATGACATTAAAATTAGAGCTATTAATTAA 798
              || || | |   | |   ||  | ||  | | |||||| | |   |  || | |
Db        161 TATTTTTAAACAATTTTTTTTGAAATAAATAAAAAAAATAAAATAACATAAAATATTATT 220
```



```
RESULT 15
AG346840/c
LOCUS       AG346840               1696 bp    DNA     linear   GSS 02-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-142C12.T7, genomic survey
            sequence.
ACCESSION   AG346840
VERSION     AG346840.1  GI:47920150
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1696)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1696
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-142C12.T7"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             4.3%;  Score 88;  DB 9;  Length 1696;
  Best Local Similarity   47.3%;  Pred. No. 5.2e-06;
  Matches  464;  Conservative    0;  Mismatches  501;  Indels   16;  Gaps    6;
```



```
Qy       1135 AAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGT 1194
              ||| ||| |||  ||  | |  | | ||  |||||| ||||||  ||| || |
Db        879 TAATAAATAAATAAATATATAATATAAACA-TATAAAAAATATATAAAAAAATATACAAA 821

Qy       1195 AAAAAGAAAGAGAAAGAAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTGATAG 1254
              | ||  ||  |  || | || |   |||  ||      |   | |||| || ||   ||
Db        820 ATAATAAATAAATAATATAAATAAATAATATAAAAAAATTAAAAAATAAATTATAATTAA 761

Qy       1255 ATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTTATATTATAAAAATTTACT 1314
              |  || || |  | |  ||| || |        |   ||||     | ||   |
Db        760 AAATA-AACATAAATACAAAATAATATAAAATAAAAAAACATATAAATATATAAAAAATA 702

Qy       1315 ATTTTGAGAAAGTTTTGAAATGTATAGAATTGAGTGGGACATCCATAAAAGGAAAGTGTA 1374
              |      | | |  | | ||||  ||| || | |     | |   | | ||  ||| | ||
Db        701 AAAAATAAATAAATATAAAATAAATATAAATAATAAAAATAAAAAAACAATAAAAATATA 642

Qy       1375 TAGAATTAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTTG 1434
              || || ||||    | |      |   | |    ||| | |    ||     |   ||  |
Db        641 TATAAATAAAAAACAAATCACAAAATAAAAAAATATAAAAATATAAAAAATAAAATAATA 582

Qy       1435 ATTTCATAAGATTATAAATCTATGTTATAATGATAATATAATTTT----AAAAATAATAC 1490
              |    | || |||| |||| |||   || |  |||||| ||| ||    ||||| || |
Db        581 AAAAAACAA-ATTAAAAATATATAAAATTAATATAATAAAATCTTTAAAAAAAAAAAAAA 523

Qy       1491 TATATTAATTCTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACTGAGTAATATG 1550
              ||  ||||  | |||| |  |  |       | ||||||  |  | | | |  || ||
Db        522 AATTATAATAATAATTAATAAAAAAAAAAAAAATAATAATTAAAATAAAATAAAAAAAATA 463

Qy       1551 AATAAATCAGTTATCTGAAAAGCAAATAATATCTTTGTAAAACAGCGTTCGGTCAAATGG 1610
              || || | |   || |  |||  ||| |  |  |||  |||| |          |||
Db        462 AAGAATTAAAAAATATACAAAAAAAAAATAAAATTTAAAAAATATTCAAAAAAAAAAAAA 403

Qy       1611 GAAGTTCATGTGTATTCAATAGTTTTAATATAAAAGTAAATTTTAAATTAATTGTTATTT 1670
              ||   |    |  || |     | || ||||  ||||   |||  ||     | |
Db        402 AAATAAAAATCAACTAAAACAAAAAAACTAAAAAAAAAAAATCAAAAAAAAAAAAAAAAATC 343

Qy       1671 TTGTTTCAGAAATTTAAAATA 1691
              || | ||| ||| | |||| |
Db        342 TTATCTCAAAAAATAAAAAAA 322
```

Search completed: December 7, 2004, 05:42:08
Job time : 7749.85 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 16:35:35 ; Search time 3914.95 Seconds
(without alignments)
10279.455 Million cell updates/sec

Title: US-09-806-197-2
Perfect score: 851
Sequence: 1 ctgaaaaggaagttcatcga.....aaaaagttgcaattctgtca 851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	851	100.0	851	6	E40088	E40088 Plant promo
c	2	123.6	14.5	739	6	E09663	E09663 cDNA encodi
c	3	123.6	14.5	739	6	E40092	E40092 Plant promo
c	4	123.6	14.5	739	6	I20016	I20016 Sequence 3
c	5	123.6	14.5	739	8	D88388	D88388 Daucus caro
c	6	122.6	14.4	774	8	AB127961	AB127961 Daucus ca
c	7	90	10.6	727	8	DCU47087	U47087 Daucus caro
c	8	88.6	10.4	2048	6	E40089	E40089 Plant promo
c	9	88.6	10.4	2048	6	E40090	E40090 Plant promo
c	10	88.6	10.4	2052	6	E40087	E40087 Plant promo
c	11	88.6	10.4	2052	6	E40093	E40093 Plant promo
c	12	88.6	10.4	2056	6	E40091	E40091 Plant promo
c	13	77.8	9.1	7218	6	I66494	I66494 Sequence 14
	14	76.4	9.0	8005	8	DCA18706	Y18706 Daucus caro
c	15	74.8	8.8	2215	8	DCU56392	U56392 Daucus caro
c	16	74.8	8.8	2215	8	S83359	S83359 iEP4=42.6-k
c	17	73.2	8.6	4886	8	DARGCHS2	D16255 Carrot gCHS
c	18	71	8.3	2042	6	AR076817	AR076817 Sequence
c	19	71	8.3	2042	6	E15125	E15125 Promoter. 7
c	20	64.6	7.6	281723	3	PFA929359	AL929359 Plasmodiu
	21	63.6	7.5	836	6	AR076819	AR076819 Sequence
	22	63.6	7.5	836	6	E15127	E15127 Terminator.
	23	63.2	7.4	46548	3	AC006622	AC006622 Caenorhab
c	24	62	7.3	2831	6	BD188672	BD188672 Promotor
c	25	62	7.3	2865	6	BD188674	BD188674 Promotor
	26	61	7.2	44352	3	AF098501	AF098501 Caenorhab
	27	60	7.1	110000	3	AC116984_2	Continuation (3 of
	28	59	6.9	110000	8	CR382131_03	Continuation (4 of
	29	58.8	6.9	110000	3	AC116305_0	AC116305 Dictyoste
c	30	58.4	6.9	16924	3	AF335994	AF335994 Lepidopso
c	31	58.4	6.9	152339	5	BX663505	BX663505 Zebrafish
c	32	58.4	6.9	273694	2	CR376734	CR376734 Danio rer
	33	58	6.8	246611	2	AC111404	AC111404 Rattus no
c	34	57.4	6.7	256879	3	AC116982	AC116982 Dictyoste
c	35	56.6	6.7	132434	5	BX005134	BX005134 Zebrafish
c	36	56.6	6.7	155790	2	BX927291	BX927291 Danio rer
c	37	56.6	6.7	173019	5	BX321884	BX321884 Zebrafish
	38	56.6	6.7	187500	2	CR450798	CR450798 Danio rer
c	39	56.6	6.7	250029	3	AE014820	AE014820 Plasmodiu
c	40	56.2	6.6	110000	3	AC116957_2	Continuation (3 of
	41	55.8	6.6	1141	6	AX083744	AX083744 Sequence
	42	55.8	6.6	341050	3	PFA929357	AL929357 Plasmodiu
c	43	55.4	6.5	53932	2	AC023371	AC023371 Homo sapi
	44	55.4	6.5	185409	9	AC018758	AC018758 Homo sapi
c	45	55.2	6.5	349980	6	AX344550	AX344550 Sequence

ALIGNMENTS

RESULT 1
E40088
LOCUS E40088 851 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40088
VERSION E40088.1 GI:18627204
KEYWORDS JP 2000166577-A/2.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 851)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 2 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000166577-A/2
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC C12N15/00,
PC C12N5/00,(C12N5/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT terminator (1). .(851).
FEATURES Location/Qualifiers
source 1. .851
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 100.0%; Score 851; DB 6; Length 851;
Best Local Similarity 100.0%; Pred. No. 3.2e-152;
Matches 851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGAAAAGGAAGTTCATCGATCTATCAGCAAAATTAGAGAACTTGTGAGGTCACAGAAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGAAAAGGAAGTTCATCGATCTATCAGCAAAATTAGAGAACTTGTGAGGTCACAGAAGT 60
```

```
Qy          61 CTGAAGGACTAGCGGAACCTGAAACTGGGTCTCAGAAGAGGATCACCTACGAGCAAGTGA 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CTGAAGGACTAGCGGAACCTGAAACTGGGTCTCAGAAGAGGATCACCTACGAGCAAGTGA 120

Qy         121 AGAAAATGGCAACTTTATTTGATGACTTGTTGATATTTATTGAGAATTACAACTTTGCAG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AGAAAATGGCAACTTTATTTGATGACTTGTTGATATTTATTGAGAATTACAACTTTGCAG 180

Qy         181 AAAAGCCAACTCTGCGGTTTCAGGTTCTGGAATTAATTAAGCTTTTACATCACTATGGAA 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AAAAGCCAACTCTGCGGTTTCAGGTTCTGGAATTAATTAAGCTTTTACATCACTATGGAA 240

Qy         241 GTGATACTATTCGAAGCGGAGTGGAGGAAGAACTTGAGTACGTGAATGAGAAAAATTCAG 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GTGATACTATTCGAAGCGGAGTGGAGGAAGAACTTGAGTACGTGAATGAGAAAAATTCAG 300

Qy         301 CAACACAGTACAAGAAAGCTCTGGAAGTAATGTTGAGAGTATGCAATAAGGAGAATACGG 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CAACACAGTACAAGAAAGCTCTGGAAGTAATGTTGAGAGTATGCAATAAGGAGAATACGG 360

Qy         361 GGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCT 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCT 420

Qy         421 ATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGA 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 ATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGA 480

Qy         481 AGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTT 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTT 540

Qy         541 GGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCA 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCA 600

Qy         601 CGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAA 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAA 660

Qy         661 AAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAA 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAA 720

Qy         721 ATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATA 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 ATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATA 780

Qy         781 CATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTG 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTG 840

Qy         841 CAATTCTGTCA 851
               ||||||||||||
Db         841 CAATTCTGTCA 851
```

```
RESULT 2
E09663/c
LOCUS       E09663                  739 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  cDNA encoding novel protein of carrot root.
ACCESSION   E09663
VERSION     E09663.1  GI:22026290
KEYWORDS    JP 1995188288-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Kubo,M. and Oita,K.
  TITLE     PROTEIN, SPECIFICALLY EXPRESSIBLE IN ROOT PART OF CARROT AND HAVING
            16KD MOLECULAR WEIGHT, ITS GENE AND PLASMID CONTAINING THE SAME
            GENE
  JOURNAL   Patent: JP 1995188288-A 1 25-JUL-1995;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L. (carrot)
            PN   JP 1995188288-A/1
            PD   25-JUL-1995
            PF   24-DEC-1993 JP 1993327943
            PI   KUBO MIYOSHI, OITA KENJI
            PC   C07K14/415,C12N15/09//A01H1/00,C12P21/02;
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   Feature is identified by experimental;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .739
            FT                   /organism='Daucus carota L.'
            FT                   /strain='Kurodagosun'
            FT                   /tissue_type='root'
            FT                   /clone='pC16'
            FT   3'UTR           479. .717
            FT   polyA_site      718. .739
            FT   CDS             14. .478
            FT                   /product='novel protein of carrot root'.
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="unidentified"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32644"
ORIGIN
```

```
  Query Match             14.5%;  Score 123.6;  DB 6;  Length 739;
  Best Local Similarity   96.9%;  Pred. No. 1e-13;
  Matches  126;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         722 TGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATAC 781
               ||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db         718 TGAGAATGAATATGATTATTGGTTTAATAATAGTGCATATTATTTAAAAAGATCGCATAC 659

Qy         782 ATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGC 841
               |||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db         658 ATTACCAGCCAGATGAAGTTATTCATCACAAGTCACAACAAAGTACAAAGAGAAAGTTGC 599

Qy         842 AATTCTGTCA 851
               ||||||||||
Db         598 AATTCTGTCA 589
```

```
RESULT 3
E40092/c
LOCUS       E40092                  739 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40092
VERSION     E40092.1  GI:18627208
KEYWORDS    JP 2000166577-A/6.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 6 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/6
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
```

```
           FT   source          1. .739
           FT                   /organism='Daucus carota L.'.
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             14.5%;  Score 123.6;  DB 6;  Length 739;
  Best Local Similarity   96.9%;  Pred. No. 1e-13;
  Matches  126;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        722 TGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATAC 781
              |||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        718 TGAGAATGAATATGATTATTGGTTTAATAATAGTGCATATTATTTAAAAAGATCGCATAC 659

Qy        782 ATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGC 841
              |||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        658 ATTACCAGCCAGATGAAGTTATTCATCACAAGTCACAACAAAGTACAAAGAGAAAGTTGC 599

Qy        842 AATTCTGTCA 851
              ||||||||||
Db        598 AATTCTGTCA 589


RESULT 4
I20016/c
LOCUS       I20016                  739 bp    DNA     linear   PAT 07-OCT-1996
DEFINITION  Sequence 3 from patent US 5512484.
ACCESSION   I20016
VERSION     I20016.1  GI:1600371
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Yamamoto,M. and Oheda,K.
  TITLE     Carrot 16 KD protein, gene coding for said protein and plasmid
            containing said gene
  JOURNAL   Patent: US 5512484-A 3 30-APR-1996;
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             14.5%;  Score 123.6;  DB 6;  Length 739;
  Best Local Similarity   96.9%;  Pred. No. 1e-13;
  Matches  126;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        722 TGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATAC 781
              |||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        718 TGAGAATGAATATGATTATTGGTTTAATAATAGTGCATATTATTTAAAAAGATCGCATAC 659

Qy        782 ATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGC 841
              |||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        658 ATTACCAGCCAGATGAAGTTATTCATCACAAGTCACAACAAAGTACAAAGAGAAAGTTGC 599

Qy        842 AATTCTGTCA 851
              ||||||||||
Db        598 AATTCTGTCA 589


RESULT 5
D88388/c
LOCUS       D88388                  739 bp    mRNA    linear   PLN 07-FEB-1999
DEFINITION  Daucus carota mRNA for cr16, complete cds.
ACCESSION   D88388
VERSION     D88388.1  GI:1663521
KEYWORDS    cr16.
SOURCE      Daucus carota (carrot)
  ORGANISM  Daucus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Apiales; Apiaceae; Apioideae; Scandiceae;
            Daucinae; Daucus.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Yamamoto,M., Torikai,S. and Oeda,K.
  TITLE     A major root protein of carrot with high homology to intracellular
            pathogen-related(PR) proteins and pollen allergens
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 739)
  AUTHORS   Torikai,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-OCT-1996) Satomi Torikai, Sumitomo Chemical Co.,
            Ltds., Biotechnology Laboratory; 4-2-1,Takatsukasa, Takarazuka,
            Hyogo 665, Japan (E-mail:torikai@bio.sumitomo-chem.co.jp,
            Tel:0797-74-2059, Fax:0797-74-2133)
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Daucus carota"
                     /mol_type="mRNA"
                     /db_xref="taxon:4039"
                     /tissue_type="root"
     gene            1. .739
                     /gene="CR16"
     CDS             14. .478
                     /gene="CR16"
                     /note="A major root protein"
                     /codon_start=1
                     /product="cr16"
                     /protein_id="BAA13604.1"
                     /db_xref="GI:1663522"
                     /translation="MGAQSHSLEITSSVSAEKIFSGIVLDVDTVIPKAAPGAYKSVDV
                     KGDGGAGTVRIITLPEGSPITSMTVRTDAVNKEALTYDSTVIDGDILLEFIESIETHM
                     VVVPTADGGSITKTTAIFHTKGDAVVPEENIKFADAQNTALFKAIEAYLIAN"
ORIGIN

  Query Match             14.5%;  Score 123.6;  DB 8;  Length 739;
  Best Local Similarity   96.9%;  Pred. No. 1e-13;
  Matches  126;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        722 TGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATAC 781
              |||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        718 TGAGAATGAATATGATTATTGGTTTAATAATAGTGCATATTATTTAAAAAGATCGCATAC 659

Qy        782 ATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGC 841
              |||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        658 ATTACCAGCCAGATGAAGTTATTCATCACAAGTCACAACAAAGTACAAAGAGAAAGTTGC 599

Qy        842 AATTCTGTCA 851
              ||||||||||
Db        598 AATTCTGTCA 589


RESULT 6
AB127961/c
LOCUS       AB127961                774 bp    mRNA    linear   PLN 26-MAY-2004
DEFINITION  Daucus carota DcPRP mRNA for pathogenesis-related protein, complete
            cds.
ACCESSION   AB127961
VERSION     AB127961.1  GI:39104473
KEYWORDS    .
SOURCE      Daucus carota (carrot)
  ORGANISM  Daucus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Apiales; Apiaceae; Apioideae; Scandiceae;
            Daucinae; Daucus.
REFERENCE   1
  AUTHORS   Sano,T., Nishimoto,M., Saburi,W., Kimura,A., Yasuda,H.,
            Uchibatake,M., Ohwada,T. and Masuda,H.
  TITLE     Isolation and Characterization of cDNA Encoding P-19.5 Protein
            Accumulated Preferentially at Early Stage of Carrot Somatic
            Embryogenesis
```

```
  JOURNAL   Plant Sci. (2004) In press
REFERENCE   2  (bases 1 to 774)
  AUTHORS   Sano,T. and Masuda,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2003) Takuma Sano, Obihiro University of
            Agriculture and Veterinary Medicine, Laboratory of Biological
            Function; Nishi 2-11, Inada-Cho, Obihiro, Hokkaido 080-8555, Japan
            (E-mail:anos@octv.ne.jp, URL:http://www.obihiro.ac.jp/
            ~kinoukaihatsu/, Tel:81-155-49-5115(ex.5553), Fax:81-155-49-5577)
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Daucus carota"
                     /mol_type="mRNA"
                     /db_xref="taxon:4039"
                     /note="pathogenesis-related protein accumulated in all
                     stages of somatic embryogenesis and plantlet, identical to
                     cr16"
     gene            1. .774
                     /gene="DcPRP"
     CDS             72. .536
                     /gene="DcPRP"
                     /codon_start=1
                     /product="pathogenesis-related protein"
                     /protein_id="BAD04049.1"
                     /db_xref="GI:39104474"
                     /translation="MGAQSHSLEITSSVSAEKIFSGIVLDVDTVIPKAAPGAYKSVDV
                     KGDGGAGTVRIITLPEGSPITSMTVRTDAVNKEALTYDSTVIDGDILLEFIESIETHM
                     VVVPTADGGSITKTTAIFHTKGDAVVPEENIKFADAQNTALFKAIEAYLIAN"
ORIGIN

  Query Match             14.4%;  Score 122.6;  DB 8;  Length 774;
  Best Local Similarity   96.9%;  Pred. No. 1.6e-13;
  Matches  125;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        723 GAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACA 782
              ||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        774 GAGAATGAATATGATTATTGGTTTAATAATAGTGCATATTATTTAAAAAGATCGCATACA 715

Qy        783 TTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGCA 842
              ||||||||||||||||||||||||||||| |||||||||||||||||| |||||||||
Db        714 TTACCAGCCAGATGAAGTTATTCATCACAAGTCACAACAAAGTACAAAGAGAAAGTTGCA 655

Qy        843 ATTCTGTCA 851
              |||||||||
Db        654 ATTCTGTCA 646


RESULT 7
DCU47087/c
LOCUS       DCU47087                727 bp    mRNA    linear   PLN 30-MAY-1996
DEFINITION  Daucus carota pathogenesis-related protein mRNA, somatic embryo
            clone Gea20, partial cds.
ACCESSION   U47087
VERSION     U47087.1  GI:1276955
KEYWORDS    .
SOURCE      Daucus carota (carrot)
  ORGANISM  Daucus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Apiales; Apiaceae; Apioideae; Scandiceae;
            Daucinae; Daucus.
REFERENCE   1  (bases 1 to 727)
  AUTHORS   Lin,X., Hwang,G.-J. and Zimmerman,J.L.
  TITLE     Isolation and characterization of a diverse set of genes from
            carrot somatic embryos
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 727)
  AUTHORS   Lin,X., Hwang,G.-J. and Zimmerman,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-1996) J. Lynn Zimmerman, Biological Sciences,
            University of Maryland Baltimore County, 5401 Wilkens Ave,
            Baltimore, MD 21228, USA
FEATURES             Location/Qualifiers
     source          1. .727
                     /organism="Daucus carota"
                     /mol_type="mRNA"
                     /strain="Danver Half-long"
                     /db_xref="taxon:4039"
                     /clone="Gea20"
                     /tissue_type="somatic embryos at the globular stage"
     CDS             <1. .507
                     /codon_start=1
                     /product="pathogenesis-related protein"
                     /protein_id="AAB01092.1"
                     /db_xref="GI:1335877"
                     /translation="LIFLLILSNNILNIMGAQSHSLEITSSVSAEKIFXXIVLDVDTV
                     IPKAAPGAYKSVDVKGDGGAGTVRIITLPEGSPITSMTVRTDAVNKEALTYDSTVIDG
                     DILLEFIESIETHMVVVPTADGGSITKTTAIFHTKGDAVVPEENIKFADAQNTALFKA
                     IEAYLIAN"
ORIGIN

  Query Match             10.6%;  Score 90;  DB 8;  Length 727;
  Best Local Similarity   93.7%;  Pred. No. 2.6e-07;
  Matches  104;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        742 GGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTT 801
              |||||||||||||||| || |||||||||||| | ||||||||||||||||||||||||
Db        727 GGTTTAATAATAGTGCATATTATTTAAAAAGAACNCATACATTACCAGCCAGATGAAGTT 668

Qy        802 ATTCATCACAA-CTCACAACAAAGTACAAAGAAAAAGTTGCAATTCTGTCA 851
              |||||||||||  |||||||||||||||||| ||||||||||||||||||
Db        667 ATTCATCACAAGTTCACAACAAAGTACAAAGAGAAAGTTGCAATTCTGTCA 617


RESULT 8
E40089/c
LOCUS       E40089                  2048 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40089
VERSION     E40089.1  GI:18627205
KEYWORDS    JP 2000166577-A/3.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2048)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 3 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/3
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2048).
FEATURES             Location/Qualifiers
     source          1. .2048
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             10.4%;  Score 88.6;  DB 6;  Length 2048;
  Best Local Similarity   59.6%;  Pred. No. 3.9e-07;
  Matches  221;  Conservative    0;  Mismatches  119;  Indels   31;  Gaps    3;

Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
              ||| |  |||||  |  || |||| |||  ||         |||          |  |  ||
Db        342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283
```

```
Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
              |  || | |    | ||||||| |     |||| | | |||| ||||||||||||| ||
Db        282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
              |||| ||||| |||  || || |||||||||||||| ||| ||||||          | |
Db        222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
              ||||| |||||||||||||||||    |||||||| |    |
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA------------------ 131

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
                 || ||| |  || |  ||| |||||| || ||| ||||| |||| |||||| |||||
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
              |   |||  ||||  | |||| |    | |||||   ||  |        |||  ||
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
              | | || |||
Db         13 TAGGGCACACA 3


RESULT 9
E40090/c
LOCUS       E40090                  2048 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40090
VERSION     E40090.1  GI:18627206
KEYWORDS    JP 2000166577-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2048)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 4 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/4
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2048).
FEATURES             Location/Qualifiers
     source          1. .2048
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             10.4%;  Score 88.6;  DB 6;  Length 2048;
  Best Local Similarity   59.6%;  Pred. No. 3.9e-07;
  Matches  221;  Conservative    0;  Mismatches  119;  Indels   31;  Gaps    3;

Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
              ||| |  |||||  |  || |||| |||  ||          |||        |  |  ||
Db        342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
              |  || | |    | ||||||| |     |||| | | |||| ||||||||||||| ||
Db        282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
              |||| ||||| |||  || || |||||||||||||| ||| ||||||          | |
Db        222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
              ||||| |||||||||||||||||    |||||||| |    |
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA------------------ 131

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
                 || ||| |  || |  ||| |||||| || ||| ||||| |||| |||||| |||||
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
              |   |||  ||||  | |||| |    | |||||   ||  |        |||  ||
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
              | | || |||
Db         13 TAGGGCACACA 3


RESULT 10
E40087/c
LOCUS       E40087                  2052 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40087
VERSION     E40087.1  GI:18627203
KEYWORDS    JP 2000166577-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2052)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 1 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/1
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2052).
FEATURES             Location/Qualifiers
     source          1. .2052
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             10.4%;  Score 88.6;  DB 6;  Length 2052;
  Best Local Similarity   59.6%;  Pred. No. 3.9e-07;
  Matches  221;  Conservative    0;  Mismatches  119;  Indels   31;  Gaps    3;

Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
              ||| |  |||||  |  || |||| |||  ||          |||        |  |  ||
Db        342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
              |  || | |    | ||||||| |     |||| | | |||| ||||||||||||| ||
Db        282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
              |||| ||||| |||  || || |||||||||||||| ||| ||||||          | |
Db        222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
              ||||| |||||||||||||||||    |||||||| |    |
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA------------------ 131
```

```
Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
Db         13 TAGGGCACACA 3


RESULT 11
E40093/c
LOCUS       E40093                2052 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40093
VERSION     E40093.1  GI:18627209
KEYWORDS    JP 2000166577-A/7.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2052)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 7 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/7
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2052).
FEATURES                Location/Qualifiers
     source             1. .2052
                        /organism="unidentified"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32644"
ORIGIN

  Query Match             10.4%;  Score 88.6;  DB 6;  Length 2052;
  Best Local Similarity   59.6%;  Pred. No. 3.9e-07;
  Matches  221;  Conservative    0;  Mismatches  119;  Indels   31;  Gaps    3;

Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
Db        342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
Db        282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
Db        222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA----------------- 131

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
Db         13 TAGGGCACACA 3


RESULT 12
E40091/c
LOCUS       E40091                2056 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40091
VERSION     E40091.1  GI:18627207
KEYWORDS    JP 2000166577-A/5.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2056)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 5 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/5
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2056).
FEATURES                Location/Qualifiers
     source             1. .2056
                        /organism="unidentified"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32644"
ORIGIN

  Query Match             10.4%;  Score 88.6;  DB 6;  Length 2056;
  Best Local Similarity   59.6%;  Pred. No. 3.9e-07;
  Matches  221;  Conservative    0;  Mismatches  119;  Indels   31;  Gaps    3;

Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
Db        342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
Db        282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
Db        222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA----------------- 131

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
Db         13 TAGGGCACACA 3
```

```
RESULT 13
I66494/c
LOCUS       I66494                 7218 bp    DNA     linear   PAT 28-DEC-1997
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   I66494
VERSION     I66494.1  GI:2724471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 7218)
  AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
  TITLE     Recombinant fowlpox virus
  JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .7218
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             9.1%;  Score 77.8;  DB 6;  Length 7218;
  Best Local Similarity   9.6%;  Pred. No. 3.4e-05;
  Matches   40;  Conservative  219;  Mismatches  156;  Indels    0;  Gaps    0;

Qy        155 ATTTATTGAGAATTACAACTTTGCAGAAAAGCCAACTCTGCGGTTTCAGGTTCTGGAATT 214
              || |||| ||  | | ||    |||   | ||          |  |    | || || |
Db       1521 ATCTATTCAGTTTCAAAAAACGGCATGTAGGCATCACTGTAATTACCTATCTATGCAAGT 1462

Qy        215 AATTAAGCTTTTACATCACTATGGAAGTGATACTATTCGAAGCGGAGTGGAGGAAGAACT 274
              | ||||     || |  | | ||| |  :: :  :   :::: :::: ::::::::::
Db       1461 AGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1402

Qy        275 TGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGTAATGTT 334
              ::: : : ::: ::::::::   :: :: : :: : :::::::    ::::: :: :
Db       1401 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1342

Qy        335 GAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGA 394
              ::::: : : :: :::::::: : ::::: : :  :::: :       : :: : :: :::
Db       1341 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1282

Qy        395 AAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAG 454
              ::::: :::::::::: :::: :    : ::: :: ::::::  :::: ::   ::: ::
Db       1281 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1222

Qy        455 CTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGT 514
                 :::::::: ::      :   :::: :: :::: :::: ::::  :: ::: ::::
Db       1221 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1162

Qy        515 AATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAG 569
              :: ::  : :::  :  :::: :   :::::::::: ::::  : :::::::::
Db       1161 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1107


RESULT 14
DCA18706
LOCUS       DCA18706               8005 bp    DNA     linear   PLN 10-FEB-1999
DEFINITION  Daucus carota Inv*Dc5 gene.
ACCESSION   Y18706
VERSION     Y18706.1  GI:4454114
KEYWORDS    beta-fructofuranosidase; Inv*Dc5 gene; invertase; isoform.
SOURCE      Daucus carota (carrot)
  ORGANISM  Daucus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Apiales; Apiaceae; Apioideae; Scandiceae;
            Daucinae; Daucus.
REFERENCE   1
  AUTHORS   Sturm,A.
  TITLE     Molecular characterisation and functional analysis of
            sucrose-cleaving enzymes in carrot (Daucus carota L.)
  JOURNAL   J. Exp. Bot. 47, 1187-1192 (1996)
REFERENCE   2  (bases 1 to 8005)
  AUTHORS   Sturm,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) A. Sturm, Friedrich Miescher-Institut,
            Maulbeerstr. 66, CH-4058 Basel, SWITZERLAND
COMMENT     Related sequence X67163.
FEATURES             Location/Qualifiers
     source          1. .8005
                     /organism="Daucus carota"
                     /mol_type="genomic DNA"
                     /cultivar="Nantaise"
                     /db_xref="taxon:4039"
     gene            1. .8005
                     /gene="Inv*Dc5"
     5'UTR           1. .2666
                     /gene="Inv*Dc5"
     CDS             join(2767. .3160,3724. .3732,4656. .5515,5605. .5766,
                     5869. .6104,6190. .6277,6384. .6587)
                     /gene="Inv*Dc5"
                     /EC_number="3.2.1.26"
                     /note="soluble acid invertase"
                     /codon_start=1
                     /product="beta-fructofuranosidase, isoform II"
                     /protein_id="CAA77266.1"
                     /db_xref="GI:4454115"
                     /db_xref="GOA:Q42722"
                     /db_xref="TrEMBL:Q42722"
                     /translation="MEHPITISHYTPLPDGEHSPSLTTTNTAEQSSRRRSLTFVLLFS
                     SILAACLVMGTMVLFPNSGNEAVEKSTVVPEETVEVAPRGVAEGVSMKSFRRPALNAE
                     PPANFPWNSNVLSWQRSSFHFQPNQNWMNDPNGPLFYKGWYHLFYQYNPDGAIWGNKI
                     VWGHAVSSDLIHWKHLPVAMVTDHWYDVNGVWTGSATILPDGQIVMLYTGSTNESVQV
                     QNLAYPADPSDPLLIEWVKYPGNPVLVPPPGIDFKDFRDPTTAWRTPEGKWRLIIGSK
                     LNKTGISLVYDTVDFKNFTLLDGVLHAVHGTGMWECVDFYPVSKFGENGLDTSFDGVG
                     VKHVMKASLDDDRNDYYAIGTYDPVSGKWVPDNPELDVGIGLRYDYGIYYASKTFYDS
                     NKKRRVLWSWIKETDSEISDVRKGWASVQGIPRTILFDPKTGSNLLQWPVEEVNKLRL
                     NKTVFENVEINTGAVLPLEIGSGSQLDITAEFEVDKESLERVQETNEVYDCKNNGGSS
                     GRGALGPFGLLILADKDLSEQTPVYFYIAKGSGGNLRTFFCADHSRSSKAVDVDKEIY
                     GSVVPVLRGEKLTMRILVDHSIVESFSQGGRTCITSRVYPTKAIYNNAKVFLFNNATE
                     ARIIASLNIWQMNTAQRQTHFADLVI"
     exon            2767. .3160
                     /gene="Inv*Dc5"
                     /number=1
     intron          3161. .3723
                     /gene="Inv*Dc5"
                     /number=1
     exon            3724. .3732
                     /gene="Inv*Dc5"
                     /number=2
     intron          3733. .4655
                     /gene="Inv*Dc5"
                     /number=2
     exon            4656. .5515
                     /gene="Inv*Dc5"
                     /number=3
     intron          5516. .5604
                     /gene="Inv*Dc5"
                     /number=3
     exon            5605. .5766
                     /gene="Inv*Dc5"
                     /number=4
     intron          5767. .5868
                     /gene="Inv*Dc5"
                     /number=4
     exon            5869. .6104
                     /gene="Inv*Dc5"
                     /number=5
     intron          6105. .6189
                     /gene="Inv*Dc5"
                     /number=5
     exon            6190. .6277
                     /gene="Inv*Dc5"
                     /number=6
     intron          6278. .6383
                     /gene="Inv*Dc5"
```

```
                 /number=6
     exon        6384. .6587
                 /gene="Inv*Dc5"
                 /number=7
     3'UTR       6588. .8005
                 /gene="Inv*Dc5"
ORIGIN

  Query Match             9.0%;  Score 76.4;  DB 8;  Length 8005;
  Best Local Similarity   66.9%;  Pred. No. 6.2e-05;
  Matches  170;  Conservative    0;  Mismatches   51;  Indels   33;  Gaps    3;

Qy        425 ATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTA 484
              | ||| ||  || ||| | |||| |||||||||||| |||||| ||||||||| || ||
Db       1040 AGGGTCAGTCATAGGGGCCGTTTGGGTTAGCTTAAAAGAAGTGATTTCTTACTTAAATTA 1099

Qy        485 ATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAA 544
              |||| |||||||| ||| ||||||          |||| ||| ||||||||||||||
Db       1100 GAAAAGTAGAGTAGAAGTGAGAAGTAA----------ATAAATTAATAAAGTGTTTGGAA 1149

Qy        545 AAGAAATAGAAGTTGT-AAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGA 603
              |||||  ||||| ||| | ||| ||| ||||||| || ||||
Db       1150 AAGAAGCAGAAGCTGTGAGAGAGAAGCTAGCATTCTCAACTT----------------- 1191

Qy        604 CTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAA 663
              || |||     |  | ||| ||||||||||||||||||  |||  ||||    ||
Db       1192 ----TTTAAAAATGCTTCTGCTTATTTACACAAACGGGTCAAGAGAAGCAGAAGACGGAA 1247

Qy        664 GCTGGAGTTACTTC 677
              || |    ||||||
Db       1248 GCAGCTCCTACTTC 1261


RESULT 15
DCU56392/c
LOCUS       DCU56392               2215 bp    DNA     linear   PLN 01-FEB-1999
DEFINITION  Daucus carota elicitor induced glycoprotein iEP4 gene, complete
            cds.
ACCESSION   U56392
VERSION     U56392.1  GI:4204869
KEYWORDS    .
SOURCE      Daucus carota (carrot)
  ORGANISM  Daucus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Apiales; Apiaceae; Apioideae; Scandiceae;
            Daucinae; Daucus.
REFERENCE   1  (bases 1 to 2215)
  AUTHORS   Bertinetti,C.M. and Ugalde,R.A.
  TITLE     Studies on the response of carrot cells to a Sclerotinia
            sclerotiorum elicitor. Induction of expression of an extracellular
            glycoprotein mRNA
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2215)
  AUTHORS   Bertinetti,C.M. and Ugalde,R.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-1996) C.M. Bertinetti, Instituto de
            Investigaciones Bioquimicas 'Fundacion Campomar', Patricias
            Argentinas 435, Capital Federal, Buenos Aires 1405, Argentina
FEATURES             Location/Qualifiers
     source          1. .2215
                     /organism="Daucus carota"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4039"
                     /cell_line="ZA+ line"
     mRNA            join(<55. .292,365. .580,1077. .1246,1358. .1628,1797. .>2077)
     CDS             join(55. .292,365. .580,1077. .1246,1358. .1628,1797. .2077)
                     /note="glycoprotein"
                     /codon_start=1
                     /product="iEP4"
                     /protein_id="AAD11468.1"
                     /db_xref="GI:4204870"
                     /translation="MGKHSAFSVFLFSILVAHSHVSASSQTCDFPAIFNFGDANSDTG
                     AFAAWFFGNPPFFGQSYFNGSAGRVSDGRLLIDFMATDLGLPFLHPYMDSLGANFSHG
                     ANFANILSTIALPTSNIIPGVRPPRGLNPVNLDIQVAQFAQFVNRSQTQGEAFDNFMP
                     KQDYFSQALYTLDIGQIDITQEFLTNKTDDEIKAVVPGLISSLSSNIQILYSLGGRSF
                     WIHNLGPNGCLPILLTLAPVPDDQLDSAGCAKRYNDLTQYFNSELKKGVDQLRTDLPL
                     AAVTYVDVYTAKYSLYQEPAKYGFTHPLETCCGFGGRYNYGEFSLCGSTITVNGTQLT
                     VGPCENPAEYINYEGQTYTQAADQITFNKISTGELSDPPNSLKTACPKLSLPRVSDI"
ORIGIN

  Query Match             8.8%;  Score 74.8;  DB 8;  Length 2215;
  Best Local Similarity   70.4%;  Pred. No. 0.00016;
  Matches  114;  Conservative    0;  Mismatches   47;  Indels    1;  Gaps    1;

Qy        437 TGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGT 496
              |||| | |||| ||| | ||| |||||||| |||||||  || |||||| |||||| | |
Db        871 TGGGGCCGTTTGGGTAAACTTTAAAAAAGT-ACTTCTTCTTTAAAGTAAAGAAGTGAACT 813

Qy        497 AGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAG 556
              |||| | | | ||||| ||  | |||||||  |||||| |||||||| ||| |      ||
Db        812 AGAAGTAAGATGTAAATTAGAACTTATAAGCGATTAAACTGTTTGGGAAACAGCAGAAAT 753

Qy        557 TTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCT 598
                  ||| |||||||||||||| || |  | |  ||   || |
Db        752 CCTGAAACAAAAGTTAGCATTCTCAAAGTACTTCTAGCTTTT 711


Search completed: December  6, 2004, 22:47:31
Job time : 3918.95 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 15:16:44 ; Search time 470.618 Seconds
(without alignments)
9492.325 Million cell updates/sec

Title: US-09-806-197-2
Perfect score: 851
Sequence: 1 ctgaaaaggaagttcatcga.....aaaaagttgcaattctgtca 851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	851	100.0	851	3	AAA37960	Aaa37960 Carrot te
c	2	123.6	14.5	739	2	AAQ92559	Aaq92559 cDNA clon
c	3	123.6	14.5	739	3	AAA37965	Aaa37965 Plasmid #
c	4	88.6	10.4	2048	3	AAA37962	Aaa37962 Plasmid #
c	5	88.6	10.4	2048	3	AAA37963	Aaa37963 Plasmid #
c	6	88.6	10.4	2052	3	AAA37959	Aaa37959 Carrot pr
c	7	88.6	10.4	2052	3	AAA37961	Aaa37961 Carrot pr
c	8	88.6	10.4	2056	3	AAA37964	Aaa37964 Plasmid #
c	9	71	8.3	2042	2	AAV15144	Aav15144 New promo
	10	63.6	7.5	836	2	AAV15146	Aav15146 Plant gen
c	11	62	7.3	2831	10	ADC56759	Adc56759 Carrot DN
c	12	62	7.3	2865	10	ADC56761	Adc56761 Carrot DN
c	13	54.6	6.4	2114	6	ABK40080	Abk40080 Human che
	14	52.8	6.2	4045	6	ABL92321	Abl92321 Chemicall
	15	52.8	6.2	4045	6	AAD22342	Aad22342 Chemicall
	16	52.2	6.1	8056	8	ABZ10246	Abz10246 Haematopo
c	17	52	6.1	539	5	ADL38243	Adl38243 Human ova
c	18	52	6.1	539	5	ADI73110	Adi73110 Human ova
c	19	52	6.1	13712	6	ABL33531	Abl33531 Human imm
c	20	51.8	6.1	19787	6	ABL33451	Abl33451 Human imm
c	21	51.6	6.1	8676	6	ABL70452	Abl70452 Chemicall
c	22	51.6	6.1	8676	6	AAS61415	Aas61415 Human gen
c	23	49.8	5.9	9157	6	ABL33500	Abl33500 Human imm
c	24	49.6	5.8	17721	6	ABL33728	Abl33728 Human imm
c	25	49.2	5.8	113515	6	ABL34175	Abl34175 Human imm
	26	49	5.8	9789	2	AAT41852	Aat41852 cDNA enco
c	27	48.4	5.7	8404	4	AAS46500	Aas46500 Tumour su
c	28	48.4	5.7	8404	6	ABL33595	Abl33595 Human imm
c	29	48	5.6	6378	6	ABL32176	Abl32176 Human imm
c	30	48	5.6	6378	6	ABQ67027	Abq67027 Human ang
	31	47.8	5.6	14924	6	ABL32225	Abl32225 Human imm
	32	47.8	5.6	14924	6	ABL54322	Abl54322 Chemicall
c	33	47.8	5.6	19380	6	AAS61427	Aas61427 Human gen
	34	47.8	5.6	29993	10	ADB37663	Adb37663 Human che
c	35	47.6	5.6	34548	6	ABL70604	Abl70604 Chemicall
c	36	47.4	5.6	615	2	ADR01494	Adr01494 A. gossyp
c	37	47.4	5.6	7195	4	AAS45325	Aas45325 Chemicall
c	38	47.4	5.6	7195	6	ABK28166	Abk28166 DNA trans
	39	47.4	5.6	9180	6	ABL33965	Abl33965 Human imm
	40	47.2	5.5	875	3	AAA01920	Aaa01920 Human col
c	41	46.8	5.5	5689	4	AAS45384	Aas45384 Chemicall
c	42	46.8	5.5	5689	4	AAS46426	Aas46426 Tumour su
c	43	46.8	5.5	5689	6	ABK28226	Abk28226 DNA trans
	44	46.6	5.5	40862	6	ABL34073	Abl34073 Human imm
	45	46.4	5.5	5145	6	ABL32349	Abl32349 Human imm

ALIGNMENTS

RESULT 1
AAA37960
ID AAA37960 standard; DNA; 851 BP.
XX
AC AAA37960;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot terminator sequence #2.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 2; Page 69-71; 81pp; English.
XX
CC This sequence represents a carrot terminator sequence. The invention
CC relates to plant promoters and terminators from Daucus carota L. which
CC are capable of expressing a gene of interest in plants. The invention
CC also includes a chimeric gene characterized in that it comprises the
CC promoter and a desired gene linked to each other in the form capable of
CC functioning. A method of producing a transformant comprises introducing
CC the promoter, the chimeric gene or a vector comprising the promoter and a
CC desired gene or terminator sequence into a host cell. The plant promoters
CC and terminators are useful in plant breeding, for e.g. fertilities of
CC plants may be controlled by expressing, in the host cells, a sense or
CC antisense gene of a male sterility related gene such as S-locus-specific

```
CC   RNase gene
XX
SQ   Sequence 851 BP; 324 A; 111 C; 182 G; 234 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 851;  DB 3;  Length 851;
  Best Local Similarity  100.0%;  Pred. No. 6.7e-178;
  Matches  851;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGAAAAGGAAGTTCATCGATCTATCAGCAAAATTAGAGAACTTGTGAGGTCACAGAAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGAAAAGGAAGTTCATCGATCTATCAGCAAAATTAGAGAACTTGTGAGGTCACAGAAGT 60

Qy         61 CTGAAGGACTAGCGGAACCTGAAACTGGGTCTCAGAAGAGGATCACCTACGAGCAAGTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGAAGGACTAGCGGAACCTGAAACTGGGTCTCAGAAGAGGATCACCTACGAGCAAGTGA 120

Qy        121 AGAAAATGGCAACTTTATTTGATGACTTGTTGATATTTATTGAGAATTACAACTTTGCAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAAAATGGCAACTTTATTTGATGACTTGTTGATATTTATTGAGAATTACAACTTTGCAG 180

Qy        181 AAAAGCCAACTCTGCGGTTTCAGGTTCTGGAATTAATTAAGCTTTTACATCACTATGGAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAAGCCAACTCTGCGGTTTCAGGTTCTGGAATTAATTAAGCTTTTACATCACTATGGAA 240

Qy        241 GTGATACTATTCGAAGCGGAGTGGAGGAAGAACTTGAGTACGTGAATGAGAAAAATTCAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGATACTATTCGAAGCGGAGTGGAGGAAGAACTTGAGTACGTGAATGAGAAAAATTCAG 300

Qy        301 CAACACAGTACAAGAAAGCTCTGGAAGTAATGTTGAGAGTATGCAATAAGGAGAATACGG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAACACAGTACAAGAAAGCTCTGGAAGTAATGTTGAGAGTATGCAATAAGGAGAATACGG 360

Qy        361 GGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCT 420

Qy        421 ATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGA 480

Qy        481 AGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTT 540

Qy        541 GGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCA 600

Qy        601 CGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAA 660

Qy        661 AAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAA 720

Qy        721 ATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATA 780

Qy        781 CATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTG 840

Qy        841 CAATTCTGTCA 851
              ||||||||||||
Db        841 CAATTCTGTCA 851


RESULT 2
AAQ92559/c
ID   AAQ92559 standard; cDNA to mRNA; 739 BP.
XX
AC   AAQ92559;
XX
DT   02-FEB-1996  (first entry)
XX
DE   cDNA clone encoding 16kD carrot root protein.
XX
KW   Carrot; root vegetable; turnip; radish; ss.
XX
OS   Daucus carota.
XX
FH   Key             Location/Qualifiers
FT   CDS             14. .478
FT                   /*tag=  a
FT                   /product= "16kD carrot root protein"
FT   3'UTR           479. .717
FT                   /*tag=  b
FT   misc_feature    718. .739
FT                   /*tag=  c
FT                   /note= "poly-A tail"
XX
PN   EP659884-A2.
XX
PD   28-JUN-1995.
XX
PF   27-DEC-1994;   94EP-00120679.
XX
PR   24-DEC-1993;   93JP-00327943.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Yamamoto M,  Oheda K;
XX
DR   WPI; 1995-226292/30.
DR   P-PSDB; AAR75645.
XX
PT   New gene encoding a 16kD carrot protein - used to develop high protein
PT   carrots and to increase the protein content of other root vegetables.
XX
PS   Example 5; Page 11; 21pp; English.
XX
CC   This sequence represents a cDNA clone encoding a 16kD carrot root
CC   protein. This sequence was isolated after a cDNA library (prepared using
CC   RNA from carrot root tissue) was screened using the probes represented by
CC   AAQ92556-8. These probes were produced after sequencing the N-terminal
CC   amino acid sequence (see AAR75646) of the 16kD protein. The protein
CC   encoded by this sequence was found to correspond to 10% of the total
CC   amount of proteins of carrot root tissue regardless of which variety of
CC   carrot is used. This sequence can be introduced into a vector. This
CC   sequence can be used to develop high protein carrots. The gene can also
CC   be introduced and expressed in other root vegetables (such as radishes,
CC   turnips, sugar beets and burdocks) to increase the amount of protein in
CC   these vegetables
XX
SQ   Sequence 739 BP; 218 A; 157 C; 151 G; 213 T; 0 U; 0 Other;

  Query Match            14.5%;  Score 123.6;  DB 2;  Length 739;
  Best Local Similarity  96.9%;  Pred. No. 1.2e-17;
  Matches  126;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        722 TGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATAC 781
              |||||||||||||||||||||||||||||||||| || ||||||||||||||||||||||
Db        718 TGAGAATGAATATGATTATTGGTTTAATAATAGTGCATATTATTTAAAAAGATCGCATAC 659

Qy        782 ATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGC 841
              |||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        658 ATTACCAGCCAGATGAAGTTATTCATCACAAGTCACAACAAAGTACAAAGAGAAAGTTGC 599

Qy        842 AATTCTGTCA 851
              ||||||||||
Db        598 AATTCTGTCA 589


RESULT 3
```

```
AAA37965/c
ID   AAA37965 standard; DNA; 739 BP.
XX
AC   AAA37965;
XX
DT   18-AUG-2000  (first entry)
XX
DE   Plasmid #4 DNA sequence used in mutation of promoter sequence.
XX
KW   Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS   Daucus carota.
XX
PN   WO200020613-A1.
XX
PD   13-APR-2000.
XX
PF   28-SEP-1999;   99WO-JP005303.
XX
PR   02-OCT-1998;   98JP-00281124.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Nishikawa S,  Oeda K;
XX
DR   WPI; 2000-303791/26.
DR   P-PSDB; AAY88621.
XX
PT   New Plant promoters and terminators from Daucus carota L., useful in
PT   plant breeding, for e.g. controlling fertilities of plants.
XX
PS   Example 8; Page 76-78; 81pp; English.
XX
CC   This sequence represents a plasmid sequence used in a method for
CC   introducing a mutation into a carrot promoter. The invention relates to
CC   plant promoters and terminators from Daucus carota L. which are capable
CC   of expressing a gene of interest in plants. The invention also includes a
CC   chimeric gene characterized in that it comprises the promoter and a
CC   desired gene linked to each other in the form capable of functioning. A
CC   method of producing a transformant comprises introducing the promoter,
CC   the chimeric gene or a vector comprising the promoter and a desired gene
CC   or terminator sequence into a host cell. The plant promoters and
CC   terminators are useful in plant breeding, for e.g. fertilities of plants
CC   may be controlled by expressing, in the host cells, a sense or antisense
CC   gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ   Sequence 739 BP; 218 A; 157 C; 151 G; 213 T; 0 U; 0 Other;

  Query Match            14.5%;  Score 123.6;  DB 3;  Length 739;
  Best Local Similarity   96.9%;  Pred. No. 1.2e-17;
  Matches  126;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        722 TGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATAC 781
              ||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||||
Db        718 TGAGAATGAATATGATTATTGGTTTAATAATAGTGCATATTATTTAAAAAGATCGCATAC 659

Qy        782 ATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGC 841
              ||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        658 ATTACCAGCCAGATGAAGTTATTCATCACAAGTCACAACAAAGTACAAAGAGAAAGTTGC 599

Qy        842 AATTCTGTCA 851
              ||||||||||
Db        598 AATTCTGTCA 589


RESULT 4
AAA37962/c
ID   AAA37962 standard; DNA; 2048 BP.
XX
AC   AAA37962;
XX
DT   18-AUG-2000  (first entry)
XX
DE   Plasmid #1 DNA sequence used in mutation of promoter sequence.
XX
KW   Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS   Daucus carota.
XX
PN   WO200020613-A1.
XX
PD   13-APR-2000.
XX
PF   28-SEP-1999;   99WO-JP005303.
XX
PR   02-OCT-1998;   98JP-00281124.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Nishikawa S,  Oeda K;
XX
DR   WPI; 2000-303791/26.
XX
PT   New Plant promoters and terminators from Daucus carota L., useful in
PT   plant breeding, for e.g. controlling fertilities of plants.
XX
PS   Example 8; Page 71-73; 81pp; English.
XX
CC   This sequence represents a plasmid sequence used in a method for
CC   introducing a mutation into a carrot promoter. The invention relates to
CC   plant promoters and terminators from Daucus carota L. which are capable
CC   of expressing a gene of interest in plants. The invention also includes a
CC   chimeric gene characterized in that it comprises the promoter and a
CC   desired gene linked to each other in the form capable of functioning. A
CC   method of producing a transformant comprises introducing the promoter,
CC   the chimeric gene or a vector comprising the promoter and a desired gene
CC   or terminator sequence into a host cell. The plant promoters and
CC   terminators are useful in plant breeding, for e.g. fertilities of plants
CC   may be controlled by expressing, in the host cells, a sense or antisense
CC   gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ   Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

  Query Match            10.4%;  Score 88.6;  DB 3;  Length 2048;
  Best Local Similarity   59.6%;  Pred. No. 7.3e-10;
  Matches  221;  Conservative    0;  Mismatches  119;  Indels   31;  Gaps    3;

Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
              ||| |  |||||  |  || |||| |||  ||          |||       |  |  ||
Db        342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
              |  || | |     | ||||||| |      |||| | | |||| ||||||||||||| ||
Db        282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
              |||| ||||| |||  || || |||||||||||||| ||| ||||||       | |
Db        222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
              ||||| |||||||||||||||||    |||||||| |   |
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA------------------ 131

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
                || ||| |  || |  ||| |||||| || ||| ||||| |||| |||||| |||||
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
              |   |||  ||||  | |||| |  | |||||   ||  |       |||  ||
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
              | | || |||
Db         13 TAGGGCACACA 3
```

RESULT 5
AAA37963/c
ID AAA37963 standard; DNA; 2048 BP.
XX
AC AAA37963;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #2 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 73-74; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match 10.4%; Score 88.6; DB 3; Length 2048;
Best Local Similarity 59.6%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

```
Qy      345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
            ||| |  |||||  |  || |||| |||  ||         |||       |  |  ||
Db      342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy      405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
            |  || | |    | ||||||| |     |||| | | |||| ||||||||||||| ||
Db      282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy      465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
            |||| ||||| |||  || || |||||||||||||| ||| ||||||       | |
Db      222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy      525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
            ||||| ||||||||||||||||||   |||||||| |   |
Db      172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA------------------ 131

Qy      585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
              || |||  |  ||  |  ||| |||||| || ||| ||||| |||| |||||| |||||
Db      130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy      645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
```

```
Db       73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy      705 AATGACAAACA 715
            | | || |||
Db       13 TAGGGCACACA 3
```

RESULT 6
AAA37959/c
ID AAA37959 standard; DNA; 2052 BP.
XX
AC AAA37959;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #1.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 69-70; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;

Query Match 10.4%; Score 88.6; DB 3; Length 2052;
Best Local Similarity 59.6%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

```
Qy      345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
            ||| |  |||||  |  || |||| |||  ||         |||       |  |  ||
Db      342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy      405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
            |  || | |    | ||||||| |     |||| | | |||| ||||||||||||| ||
Db      282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy      465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
            |||| ||||| |||  || || |||||||||||||| ||| ||||||       | |
Db      222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy      525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
            ||||| ||||||||||||||||||   |||||||| |   |
```

```
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA----------------- 131

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
              || ||| |  || |  ||| |||||| || ||| ||||| |||| |||||| |||||
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
              |   |||  ||||  | |||| |    | |||||   ||  |        |||  ||
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
              | | || |||
Db         13 TAGGGCACACA 3
```

RESULT 7
AAA37961/c
ID AAA37961 standard; DNA; 2052 BP.
XX
AC AAA37961;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #2.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 78-79; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 736 A; 318 C; 317 G; 681 T; 0 U; 0 Other;

Query Match 10.4%; Score 88.6; DB 3; Length 2052;
Best Local Similarity 59.6%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

```
Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
              ||| |  |||||   |  || |||| |||  ||          |||         |  |  ||
Db        342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
              |  || | |    | ||||||| |      |||| | | |||| ||||||||||||| ||
Db        282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223
```

```
Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
              |||| ||||| |||   || || ||||||||||||||| ||| ||||||         | |
Db        222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
              ||||| ||||||||||||||||||   |||||||| |    |
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA----------------- 131

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
              || ||| |  || |  ||| |||||| || ||| ||||| |||| |||||| |||||
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
              |   |||  ||||  | |||| |    | |||||   ||  |        |||  ||
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
              | | || |||
Db         13 TAGGGCACACA 3
```

RESULT 8
AAA37964/c
ID AAA37964 standard; DNA; 2056 BP.
XX
AC AAA37964;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #3 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 74-76; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;

Query Match 10.4%; Score 88.6; DB 3; Length 2056;
Best Local Similarity 59.6%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

```
Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
              ||| |  |||||  |  || |||| |||  ||         |||        |  |  ||
Db        342 AATTATTAGAATTTGTTGACACGTTAAAAAATAACCCTGTTCACGTATTGTTATCCTAAA 283

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464
              |  || | |    | ||||||| |     |||| | | |||| ||||||||||||| ||
Db        282 ATTGAAAGATACATTTATGAATGATCTTTTATTGAGGCCGTTTGGGTTAGCTTAAAAGAA 223

Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
              |||| ||||| |||  || || ||||||||||||| ||| ||||||        | |
Db        222 GTGATTTCTTGCTTATAGCAAAGAAGTGGAGTAGAAATGAGAAGTAA----------AAA 173

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584
              ||||| ||||||||||||||||||   |||||||| |   |
Db        172 AGTTAATAAAGTGTTTGGAAAAGAAGCGGAAGTTGTGAGAGA----------------- 131

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644
                || ||| |   || |  ||| |||||| || ||| ||||| |||| |||||| |||||
Db        130 --GAAGTTAGTATTCGCAGCTTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704
              |   |||  ||||  | |||| |    | |||||   ||  |       |||  ||
Db         73 AAAGAAGAAGAAGTCAGAAGCAGCTTCTGCTTCTCTCAACCAAACAGGCCCTATGTGCTG 14

Qy        705 AATGACAAACA 715
              | | || |||
Db         13 TAGGGCACACA 3
```

RESULT 9
AAV15144/c
ID AAV15144 standard; DNA; 2042 BP.
XX
AC AAV15144;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX
PN EP824150-A2.
XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Torikai S, Oeda K;
XX
DR WPI; 1998-122310/12.
XX
PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
PS Claim 2; Page 15-16; 31pp; English.
XX
CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX

SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;

Query Match 8.3%; Score 71; DB 2; Length 2042;
Best Local Similarity 73.6%; Pred. No. 5.5e-06;
Matches 117; Conservative 0; Mismatches 40; Indels 2; Gaps 2;

```
Qy        428 GTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATG 487
              | || ||    ||   |||| | | | |||||||||||||| | |||| | | |||||| |
Db       1159 GGGATGAGAAAGGGTCGTTTGGTTAAACTTAAAAAAAGTGTC-TCTTGCATAAAGTAAAG 1101

Qy        488 AAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAG 547
              ||||| |  |||| | | ||||||| ||| | |||||||| |||||| ||||||| ||||
Db       1100 AAGTGCACCAGAAGTAAGAAGTAAATTAAGATTTATAAGTGATTAAATTGTTTGGGAAAG 1041

Qy        548 AAATAGAAGTTGT-AAAGAAAAGTTAGCATTTTCTACTT 585
              ||  |||||| || |||  |||||||| ||| ||   ||
Db       1040 AAGCAGAAGTCGTGAAACTAAAGTTAGTATTGTCGGTTT 1002
```

RESULT 10
AAV15146
ID AAV15146 standard; DNA; 836 BP.
XX
AC AAV15146;
XX
DT 02-JUL-1998 (first entry)
XX
DE Plant gene terminator sequence.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW terminator; ds.
XX
OS Daucus carota.
XX
PN EP824150-A2.
XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Torikai S, Oeda K;
XX
DR WPI; 1998-122310/12.
XX
PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
PS Claim 25; Page 20; 31pp; English.
XX
CC The present sequence represents a terminator isolated from a carrot
CC genomic library of cultivar Kuroda Gosun. The specification describes a
CC novel promoter, also isolated from the same library. The novel promoter
CC can be used to direct root-specific expression in plant cells. Since the
CC promoter enables expression of a desired protein in the roots of a plant,
CC it is useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
SQ Sequence 836 BP; 278 A; 122 C; 141 G; 295 T; 0 U; 0 Other;

Query Match 7.5%; Score 63.6; DB 2; Length 836;
Best Local Similarity 71.2%; Pred. No. 0.0002;
Matches 84; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

```
Qy        437 TGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGT 496
              |||| | |||| | |   |||||||||||||||| || ||| || |||  |||| || |
Db        518 TGGGGCAGTTTGGCTGGACTTAAAAAAAGTGACTTATTGCTTAAAATAAATAAGTAGATT 577
```

```
Qy        497 AGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGA 554
              | || ||| |||| | |   | ||||||||||||||||    || ||| | | || ||
Db        578 ATAAGTGAAAAGTTGATTTGGACTTATAAGTTATTAAAAGTGTTTGAATATATATTGA 635
```

RESULT 11
ADC56759/c
ID ADC56759 standard; DNA; 2831 BP.
XX
AC ADC56759;
XX
DT 18-DEC-2003 (first entry)
XX
DE Carrot DNA that encodes an expression inducer type promoter.
XX
KW carrot; expression inducer type promoter; plant; vector; transformant;
KW ds.
XX
OS Daucus carota subsp. sativus.
XX
PN JP2003000252-A.
XX
PD 07-JAN-2003.
XX
PF 31-MAY-2001; 2001JP-00164069.
XX
PR 31-MAY-2001; 2001JP-00164069.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
DR WPI; 2003-472920/45.
XX
PT An expression inducer type promoter derived from DNA of Daucus carota L.
PT var. sativa DC. of 2831 bases and its analogues having the same
PT functions.
XX
PS Claim 1; SEQ ID NO 1; 16pp; Japanese.
XX
CC This invention relates to carrot DNA used as an expression inducer type
CC promoter. Specifically, it refers to promoter sequences derived from
CC plant DNA, preferably carrot, and functional mutants thereof that can be
CC used as expression inducers. Furthermore, the present invention describes
CC DNA, vectors, transformants and the process by which to prepare
CC transformants. This polynucleotide sequence is the DNA encoding the
CC carrot promoter of the invention.
XX
SQ Sequence 2831 BP; 950 A; 456 C; 395 G; 1030 T; 0 U; 0 Other;

Query Match 7.3%; Score 62; DB 10; Length 2831;
Best Local Similarity 73.0%; Pred. No. 0.00056;
Matches 111; Conservative 0; Mismatches 30; Indels 11; Gaps 2;

```
Qy        436 TTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAG 495
              | ||| | |||| |||||  |||||| || |||||||| |||  ||||| |||||| ||
Db       1944 TAGGGGCCGTTTGGGTTAATTTAAAAGAATTGACTTCTTGCTTATAGTAAAGAAGTGAAG 1885

Qy        496 TAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAA 555
               |||| ||| |||||           ||||||  ||||||||||||||||||  ||||
Db       1884 CAGAAGTGAGAAGTAGG----------TAAGTTGATAAAGTGTTTGGAAAAGAAGCAGAA 1835

Qy        556 GTTGT-AAAGAAAAGTTAGCATTTTCTACTTC 586
              | ||| | ||| ||  ||||||| ||  ||||
Db       1834 GCTGTGAGAGAGAAACTAGCATTCTCAGCTTC 1803
```

RESULT 12
ADC56761/c
ID ADC56761 standard; DNA; 2865 BP.
XX
AC ADC56761;
XX
DT 18-DEC-2003 (first entry)
XX
DE Carrot DNA encoding an expression inducer type promoter (SeqID 3).
XX
KW carrot; expression inducer type promoter; plant; vector; transformant;
KW ds.
XX
OS Daucus carota subsp. sativus.
XX
PN JP2003000252-A.
XX
PD 07-JAN-2003.
XX
PF 31-MAY-2001; 2001JP-00164069.
XX
PR 31-MAY-2001; 2001JP-00164069.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
DR WPI; 2003-472920/45.
XX
PT An expression inducer type promoter derived from DNA of Daucus carota L.
PT var. sativa DC. of 2831 bases and its analogues having the same
PT functions.
XX
PS Disclosure; SEQ ID NO 3; 16pp; Japanese.
XX
CC This invention relates to carrot DNA used as an expression inducer type
CC promoter. Specifically, it refers to promoter sequences derived from
CC plant DNA, preferably carrot, and functional mutants thereof that can be
CC used as expression inducers. Furthermore, the present invention describes
CC DNA, vectors, transformants and the process by which to prepare
CC transformants. This polynucleotide sequence is DNA encoding a carrot
CC promoter (SeqID 3) of the invention.
XX
SQ Sequence 2865 BP; 963 A; 461 C; 396 G; 1045 T; 0 U; 0 Other;

Query Match 7.3%; Score 62; DB 10; Length 2865;
Best Local Similarity 73.0%; Pred. No. 0.00057;
Matches 111; Conservative 0; Mismatches 30; Indels 11; Gaps 2;

```
Qy        436 TTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAG 495
              | ||| | |||| |||||  |||||| || |||||||| |||  ||||| |||||| ||
Db       1944 TAGGGGCCGTTTGGGTTAATTTAAAAGAATTGACTTCTTGCTTATAGTAAAGAAGTGAAG 1885

Qy        496 TAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAA 555
               |||| ||| |||||           ||||||  ||||||||||||||||||  ||||
Db       1884 CAGAAGTGAGAAGTAGG----------TAAGTTGATAAAGTGTTTGGAAAAGAAGCAGAA 1835

Qy        556 GTTGT-AAAGAAAAGTTAGCATTTTCTACTTC 586
              | ||| | ||| ||  ||||||| ||  ||||
Db       1834 GCTGTGAGAGAGAAACTAGCATTCTCAGCTTC 1803
```

RESULT 13
ABK40080/c
ID ABK40080 standard; DNA; 2114 BP.
XX
AC ABK40080;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #81 strand 2.
XX
KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200202806-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007470.

```
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-154757/20.
XX
PT   New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT   useful for detecting cytosine methylation state of genes associated with
PT   pharmacogenomics and for therapy of diseases e.g. cancer.
XX
PS   Claim 1; SEQ ID NO 162; 24pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence at least 18
CC   bases in length of a segment of the chemically pretreated DNA of genes
CC   associated with pharmacogenomics according to one of the sequences of the
CC   genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1 (NM_000497), CYP3A3
CC   (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2 (NM_001979), OCLN
CC   (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360), MRP (NM_004996,
CC   NM_019900, NM_019901, NM_019902, NM_019862, NM_019898, NM_019899) and
CC   their complementary sequences, or a sequence (S1) chosen from 87
CC   sequences and their complements. The chemical pretreatment is bisulphite
CC   treatment to convert cytosines (but not methyl-cytosines) into uracils.
CC   Also included are an oligomer (II) in particular an oligonucleotide or a
CC   peptide nucleic acid (PNA)-oligomer, comprising in each case at least one
CC   base sequence having a length of 9 nucleotides which hybridises to or is
CC   identical to a chemically pretreated DNA of genes associated with
CC   pharmacogenomics and their complements, arranged in an array for
CC   analysing diseases associated with the methylation state (CpG) and/or
CC   detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The
CC   oligomers may also be used as PCR primers. The set of 87 nucleic acids
CC   and their complements is useful for diagnosis and therapy of solid
CC   tumours and cancer. The present sequence represents one the 87 DNA
CC   sequences or its complement. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2114 BP; 504 A; 37 C; 451 G; 1122 T; 0 U; 0 Other;

  Query Match             6.4%;  Score 54.6;  DB 6;  Length 2114;
  Best Local Similarity   56.4%;  Pred. No. 0.023;
  Matches  102;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy        649 AAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATG 708
              || |  || | ||||  |  | ||| |  | ||||  || | ||||  |||| | |||
Db       1499 AAATAAAAACCAAAAAATATATTTAATAATCATAATTATATATATATAAAATAAAAAACA 1440

Qy        709 ACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAA 768
              ||||  | | |||  | ||  ||    | |     |   | ||  | |||||  |||||
Db       1439 ACAATAATACAAAAAATAAATAAATAAAATTAAAAATATCTTATTATTTATTTATTTTAA 1380

Qy        769 AAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACA 828
              ||  |   | ||| |||| | ||| || ||  ||       |  | || ||||  | ||||
Db       1379 AATAAAATCTTACTTTACTACCCAAATTAAAATACAATAAAATAATCTCAACTCACTACA 1320

Qy        829 A 829
              |
Db       1319 A 1319


RESULT 14
ABL92321
ID   ABL92321 standard; DNA; 4045 BP.
XX
AC   ABL92321;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Chemically treated DNA repair gene fragment complementary to#65.
XX
KW   DNA repair; cytosine methylation; PMS2L1; PMS2L12; PMS2L2; PMS2L3; PMS2;
KW   L4; PMS2L5; PMS2L6; MGMT; MSH2; NUDT1; TDG; INPPL1; RFC4; DDIT1L; FANCB;
KW   XRCC8; ataxia telangiectasia; aging; Bloom's syndrome; Cockayne syndrome;
KW   Nijmegen breakage syndrome; Werner syndrome; immunodeficiency;
KW   trichthiodystrophy; Fanconi's anaemia; solid tumour; cancer; ds.
XX
OS   Unidentified.
XX
PN   WO200181622-A2.
XX
PD   01-NOV-2001.
XX
PF   06-APR-2001; 2001WO-EP003972.
XX
PR   06-APR-2000; 2000DE-01019058.
PR   07-APR-2000; 2000DE-01019173.
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-034446/04.
XX
PT   New nucleic acid derived from genes associated with DNA repair, useful
PT   for diagnosis, e.g. of ataxia telangiectasia, by determination of
PT   cytosine methylation.
XX
PS   Claim 1; SEQ ID NO 130; 25pp + Sequence Listing; English.
XX
CC   The invention relates to nucleic acids containing a sequence of at least
CC   18 nucleotides of chemically treated DNA of genes associated with DNA
CC   repair, and their complements. The invention also relates to nucleic
CC   acids comprising at least 18 base pairs of the chemically pretreated DNA
CC   of genes associated with DNA repair selected from PMS2L1, PMS2L12,
CC   PMS2L2, PMS2L3, PMS2, L4, PMS2L5, PMS2L6, MGMT, MSH2, NUDT1, TDG, INPPL1,
CC   RFC4, DDIT1L, FANCB, or XRCC8. Nucleic acids of the invention and related
CC   oligomers, are useful for diagnosis of diseases associated with gene
CC   repair, specifically ataxia telangiectasia, aging, Bloom's syndrome,
CC   Cockayne syndrome, Nijmegen breakage syndrome or Werner syndrome,
CC   immunodeficiency, trichthiodystrophy, Fanconi's anaemia, solid tumours
CC   and cancer, particularly by determining status of cytosine methylation
CC   and/or by detecting single-nucleotide polymorphisms. Determination of
CC   individual methylation patterns may allow development of individualised
CC   therapies. The sequences given in records ABL92192-ABL92335 represent
CC   chemically pre-treated DNA fragments from genes associated with DNA
CC   repair, and their complements. Note: The sequence data for this patent is
CC   not represented in the specification, but is based on sequence
CC   information supplied by the European Patent Office
XX
SQ   Sequence 4045 BP; 1310 A; 64 C; 694 G; 1977 T; 0 U; 0 Other;

  Query Match             6.2%;  Score 52.8;  DB 6;  Length 4045;
  Best Local Similarity   47.5%;  Pred. No. 0.064;
  Matches  188;  Conservative    0;  Mismatches  207;  Indels    1;  Gaps    1;

Qy        374 TATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGG 433
              | | |||||  | | |||| |    | | |||   | ||||      | |||  | |
Db        858 TTTGTTTTAAAAGAAAATATATTTTGTAAAAAAAAAAAAAGAAAAAAAAGAAAAGAAAAA 917

Qy        434 AATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGG 493
              ||  |||    || |||    |  || |  | | | |  | |||  ||   |||| || |
Db        918 AAGGGGGGGAGTGTAGTGGAAATTGAGTATAATAAAATGTTAAGTGTTTTAATAAAATAA 977

Qy        494 AGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAG 553
              | || |  | | | ||  |    |    |   || ||||||| ||||  ||||  | ||
Db        978 AATAAACGTAAAATGTTTTGAGGGGAGAGTTTTTTTTTAAAGTATTTGTTAAAGTTAGAG 1037

Qy        554 AAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAA 613
              |||   ||| |   | || |      |    |   | |  ||| | |    || | |
Db       1038 AAGAAATAATGGGTATTTTGTTAAGTTAGGGTAAGAATATTTTTTAAAATTTTATAAGTG 1097
```

```
Qy        614 GTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAG-CAAAAAGCTGGAGTT 672
              | || ||    ||||| | |       | |      | ||||||  ||| | |     |
Db       1098 TAAATTTTTTTAGTTTTAAATATTTATATAAGTTTTATTGGAAGATAAATATATTTAGAGT 1157

Qy        673 ACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAAT 732
              | || | ||| | |   ||  |||| |||| |       | |||     | ||
Db       1158 AATTTTAATATGGTTTAAATTATGTAATGTTAAATAATTTTTTTAAAAAAAGATTTTAAG 1217

Qy        733 ATGATTATTGGTTTAATAATAGTGTATTTTATTTAA 768
                | | |  | || || |      ||| | ||| |
Db       1218 TAAAATGTATGGTTTATGAATAGAAATTATGTTTTA 1253


RESULT 15
AAD22342
ID   AAD22342 standard; DNA; 4045 BP.
XX
AC   AAD22342;
XX
DT   12-FEB-2002  (first entry)
XX
DE   Chemically treated human genomic DNA #32 associated with DNA adducts.
XX
KW   DNA adduct; peptide nucleic acid; PNA; cytosine methylation;
KW   gene therapy; tumour; cancer; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO200177378-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-EP004015.
XX
PR   06-APR-2000; 2000DE-01019058.
PR   07-APR-2000; 2000DE-01019173.
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-010923/01.
XX
PT   Novel nucleic acid comprising sequence of a segment of chemically
PT   pretreated DNA of genes associated with DNA adduct, useful for diagnosis
PT   and therapy of solid tumors and cancer.
XX
PS   Claim 1; Page 93-94; 97pp; English.
XX
CC   The invention relates to chemically modified DNA of genes associated with
CC   DNA adducts, oligonucleotides and/or peptide nucleic acid (PNA) oligomers
CC   for detecting cytosine methylations as well as method for ascertaining
CC   genetic and/or epigenetic parameters of genes associated with DNA
CC   adducts. Oligomers of the invention coupled to a solid phase is useful
CC   for manufacturing an arrangement of different oligomers (array) fixed to
CC   a carrier material for analysing diseases associated with the methylation
CC   state of the CpG dinucleotides of chemically pretreated DNA of genes
CC   associated with DNA adduct. They are also useful for ascertaining genetic
CC   and/or epigenetic parameters for the diagnosis and/or therapy of existing
CC   diseases or the predisposition to specific diseases by analysing cytosine
CC   methylations. Sequences of the invention are useful for diagnosis and
CC   therapy of solid tumours and cancers. They are also useful in gene
CC   therapy. The present sequence is chemically pretreated human genomic DNA
CC   associated DNA adducts
XX
SQ   Sequence 4045 BP; 1310 A; 64 C; 694 G; 1977 T; 0 U; 0 Other;

  Query Match             6.2%;  Score 52.8;  DB 6;  Length 4045;
  Best Local Similarity   47.5%;  Pred. No. 0.064;
  Matches  188;  Conservative    0;  Mismatches  207;  Indels    1;  Gaps    1;

Qy        374 TATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGG 433
              | | |||||  | | |||| |    | | |||   | ||||      | |||  | |
Db        858 TTTGTTTTAAAAGAAAATATATTTTGTAAAAAAAAAAAAAAGAAAAAAAAGAAAAGAAAAA 917

Qy        434 AATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGG 493
              ||  |||   || |||   |  || |  | | | |  | |||  ||   |||| || |
Db        918 AAGGGGGGGAGTGTAGTGGAAATTGAGTATAATAAAATGTTAAGTGTTTTAATAAAATAA 977

Qy        494 AGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAG 553
              | || |  | | | ||   |      |   |   || |||||| ||||  ||||  | ||
Db        978 AATAAACGTAAAATGTTTTGAGGGGAGAGTTTTTTTTTAAAGTATTTGTTAAAGTTAGAG 1037

Qy        554 AAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAA 613
              |||   ||| |   | || |      |      |   | |  ||| | |    || | |
Db       1038 AAGAAATAATGGGTATTTTGTTAAGTTAGGGTAAGAATATTTTTTAAAATTTTATAAGTG 1097

Qy        614 GTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAG-CAAAAAGCTGGAGTT 672
              | || ||    ||||| | |       | |      | ||||||  ||| | |     |
Db       1098 TAAATTTTTTTAGTTTTAAATATTTATATAAGTTTTATTGGAAGATAAATATATTAGAGT 1157

Qy        673 ACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAAT 732
              | || | ||| | |   ||  |||| |||| |       | |||     | ||
Db       1158 AATTTTAATATGGTTTAAATTATGTAATGTTAAATAATTTTTTTAAAAAAAGATTTTAAG 1217

Qy        733 ATGATTATTGGTTTAATAATAGTGTATTTTATTTAA 768
                | | |  | || || |      ||| | ||| |
Db       1218 TAAAATGTATGGTTTATGAATAGAAATTATGTTTTA 1253


Search completed: December  6, 2004, 19:02:28
Job time : 472.618 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:08:52 ; Search time 85.3031 Seconds
(without alignments)
7090.965 Million cell updates/sec

Title: US-09-806-197-2
Perfect score: 851
Sequence: 1 ctgaaaaggaagttcatcga.....aaaaagttgcaattctgtca 851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	123.6	14.5	739	1	US-08-363-010-3	Sequence 3, Appli
c	2	77.8	9.1	7218	1	US-08-232-463-14	Sequence 14, Appl
c	3	71	8.3	2042	2	US-08-911-434A-2	Sequence 2, Appli
	4	63.6	7.5	836	2	US-08-911-434A-5	Sequence 5, Appli
	5	55.8	6.6	1141	4	US-09-806-708B-22	Sequence 22, Appl
c	6	47.4	5.6	615	3	US-08-998-416-186	Sequence 186, App
c	7	45.2	5.3	1141	4	US-09-806-708B-22	Sequence 22, Appl
	8	44.8	5.3	1341	3	US-08-960-780-51	Sequence 51, Appl
	9	44.8	5.3	1341	3	US-09-073-898-51	Sequence 51, Appl
	10	44.8	5.3	1341	3	US-09-371-913A-2	Sequence 2, Appli
	11	44.8	5.3	1341	4	US-09-307-106-11	Sequence 11, Appl
	12	44.8	5.3	1341	4	US-09-967-805-2	Sequence 2, Appli
	13	44.8	5.3	1341	4	US-09-850-351A-51	Sequence 51, Appl
	14	44.8	5.3	4041	1	US-08-471-033-22	Sequence 22, Appl
	15	44.8	5.3	4041	2	US-08-471-044-22	Sequence 22, Appl
	16	44.8	5.3	4041	2	US-08-463-483A-22	Sequence 22, Appl
	17	44.8	5.3	4041	2	US-08-471-046A-22	Sequence 22, Appl
	18	44.8	5.3	4041	2	US-08-470-566B-22	Sequence 22, Appl
	19	44.8	5.3	4041	2	US-08-469-334-22	Sequence 22, Appl
	20	44.8	5.3	4041	3	US-09-300-529-22	Sequence 22, Appl
	21	44.8	5.3	6049	1	US-08-471-033-1	Sequence 1, Appli
	22	44.8	5.3	6049	2	US-08-471-044-1	Sequence 1, Appli
	23	44.8	5.3	6049	2	US-08-463-483A-1	Sequence 1, Appli
	24	44.8	5.3	6049	2	US-08-471-046A-1	Sequence 1, Appli
	25	44.8	5.3	6049	2	US-08-470-566B-1	Sequence 1, Appli
	26	44.8	5.3	6049	2	US-08-469-334-1	Sequence 1, Appli
	27	44.8	5.3	6049	3	US-09-300-529-1	Sequence 1, Appli
c	28	44.6	5.2	6866	4	US-10-204-708-20	Sequence 20, Appl
	29	43.6	5.1	1296	4	US-09-248-796A-2883	Sequence 2883, Ap
	30	43.4	5.1	1356	4	US-09-307-106-49	Sequence 49, Appl
	31	42.8	5.0	705	4	US-09-270-767-5061	Sequence 5061, Ap
	32	42.8	5.0	705	4	US-09-270-767-20343	Sequence 20343, A
c	33	42.8	5.0	5152	4	US-10-204-708-47	Sequence 47, Appl
	34	42.6	5.0	5152	4	US-10-204-708-73	Sequence 73, Appl
	35	42.6	5.0	319608	4	US-09-539-333D-1	Sequence 1, Appli
	36	42.6	5.0	319608	4	US-09-679-409-1	Sequence 1, Appli
c	37	42.4	5.0	1449	4	US-09-107-532A-2018	Sequence 2018, Ap
	38	42.2	5.0	4074	1	US-08-471-033-19	Sequence 19, Appl
	39	42.2	5.0	4074	2	US-08-471-044-19	Sequence 19, Appl
	40	42.2	5.0	4074	2	US-08-463-483A-19	Sequence 19, Appl
	41	42.2	5.0	4074	2	US-08-471-046A-19	Sequence 19, Appl
	42	42.2	5.0	4074	2	US-08-470-566B-19	Sequence 19, Appl
	43	42.2	5.0	4074	2	US-08-469-334-19	Sequence 19, Appl
	44	42.2	5.0	4074	3	US-09-300-529-19	Sequence 19, Appl
	45	42	4.9	98844	4	US-09-791-211-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-363-010-3/c
; Sequence 3, Application US/08363010
; Patent No. 5512484
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Mika
; APPLICANT: Oheda, Kenji
; TITLE OF INVENTION: CARROT 16 KD PROTEIN, GENE CODING FOR
; TITLE OF INVENTION: SAID PROTEIN AND PLASMID CONTAINING SAID GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,010
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 20-3628P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota
; STRAIN: Kuroda Gosun
US-08-363-010-3

Query Match 14.5%; Score 123.6; DB 1; Length 739;
Best Local Similarity 96.9%; Pred. No. 9.3e-22;

```
  Matches  126;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        722 TGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATAC 781
              ||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db        718 TGAGAATGAATATGATTATTGGTTTAATAATAGTGCATATTATTTAAAAAGATCGCATAC 659

Qy        782 ATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGC 841
              ||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        658 ATTACCAGCCAGATGAAGTTATTCATCACAAGTCACAACAAAGTACAAAGAGAAAGTTGC 599

Qy        842 AATTCTGTCA 851
              ||||||||||
Db        598 AATTCTGTCA 589


RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             9.1%;  Score 77.8;  DB 1;  Length 7218;
  Best Local Similarity   9.6%;  Pred. No. 5.9e-10;
  Matches   40;  Conservative  219;  Mismatches  156;  Indels    0;  Gaps    0;

Qy        155 ATTTATTGAGAATTACAACTTTGCAGAAAAGCCAACTCTGCGGTTTCAGGTTCTGGAATT 214
              || |||| ||  | | ||    |||   | ||           |  |    | || || |
Db       1521 ATCTATTCAGTTTCAAAAAACGGCATGTAGGCATCACTGTAATTACCTATCTATGCAAGT 1462

Qy        215 AATTAAGCTTTTACATCACTATGGAAGTGATACTATTCGAAGCGGAGTGGAGGAAGAACT 274
              | ||||     || |  | | ||| |  :: :  :   :::: :::: ::::::::::
Db       1461 AGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1402

Qy        275 TGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGTAATGTT 334
              ::: : : ::: ::::::::  :: :: : :: : :::::::      ::::: :: :
Db       1401 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1342

Qy        335 GAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGA 394
              ::::: : : :: :::::::: : ::::: : :  :::: :     : :: : :: :::
Db       1341 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1282

Qy        395 AAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAG 454
              ::::: :::::::::: :::: :    : ::: :: :::::  :::: ::   :::  ::
Db       1281 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1222

Qy        455 CTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGT 514
              :::::::: ::      :  :::: :: :::: :::: :::  :: ::: ::::
Db       1221 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1162

Qy        515 AATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAG 569
              :: ::  : :::  :  :::: :   :::::::::: ::::  : :::::::::
Db       1161 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1107


RESULT 3
US-08-911-434A-2/c
; Sequence 2, Application US/08911434A
; Patent No. 5959176
;  GENERAL INFORMATION:
;    APPLICANT:  TORIKAI, Satomi
;    APPLICANT:  OEDA, Kenji
;    TITLE OF INVENTION:  PLANT PROMOTER AND UTILIZATION THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH, LLP
;      STREET:  P.O. BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VIRGINIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  22040
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/911,434A
;      FILING DATE:  12-AUG-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Raymond C.
;      REGISTRATION NUMBER:  21,066
;      REFERENCE/DOCKET NUMBER:  2185-0199P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)205-8000
;      TELEFAX:  (703)205-8050
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2042 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Daucus carota L.
;      INDIVIDUAL ISOLATE:  Kuroda Gosun
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  1..2042
US-08-911-434A-2

  Query Match             8.3%;  Score 71;  DB 2;  Length 2042;
```

```
  Best Local Similarity   73.6%;  Pred. No. 1.9e-08;
  Matches  117;  Conservative    0;  Mismatches   40;  Indels    2;  Gaps    2;

Qy        428 GTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATG 487
              | || ||     ||   |||| | | | ||||||||||||| | |||| | | |||||| |
Db       1159 GGGATGAGAAAGGGTCGTTTGGTTAAACTTAAAAAAAGTGTC-TCTTGCATAAAGTAAAG 1101

Qy        488 AAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAG 547
              ||||| |  |||| | | ||||||| ||| | |||||||| |||||| ||||||| ||||
Db       1100 AAGTGCACCAGAAGTAAGAAGTAAATTAAGATTTATAAGTGATTAAATTGTTTGGGAAAG 1041

Qy        548 AAATAGAAGTTGT-AAAGAAAAGTTAGCATTTTCTACTT 585
              ||  |||||| || |||  |||||||| ||| ||   ||
Db       1040 AAGCAGAAGTCGTGAAACTAAAGTTAGTATTGTCGGTTT 1002


RESULT 4
US-08-911-434A-5
; Sequence 5, Application US/08911434A
; Patent No. 5959176
;  GENERAL INFORMATION:
;    APPLICANT:  TORIKAI, Satomi
;    APPLICANT:  OEDA, Kenji
;    TITLE OF INVENTION:  PLANT PROMOTER AND UTILIZATION THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH, LLP
;      STREET:  P.O. BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VIRGINIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  22040
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/911,434A
;      FILING DATE:  12-AUG-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Raymond C.
;      REGISTRATION NUMBER:  21,066
;      REFERENCE/DOCKET NUMBER:  2185-0199P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)205-8000
;      TELEFAX:  (703)205-8050
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  836 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Daucus carota L.
;      INDIVIDUAL ISOLATE:  Kuroda Gosun
;    FEATURE:
;      NAME/KEY:  terminator
;      LOCATION:  1..836
US-08-911-434A-5

  Query Match             7.5%;  Score 63.6;  DB 2;  Length 836;
  Best Local Similarity   71.2%;  Pred. No. 9.8e-07;
  Matches   84;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy        437 TGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGT 496
              |||| | |||| | |   |||||||||||||||||| || ||| || |||  |||| || |
Db        518 TGGGGCAGTTTGGCTGGACTTAAAAAAAGTGACTTATTGCTTAAAATAAATAAGTAGATT 577

Qy        497 AGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGA 554
              | || ||| ||||  | |   | ||||||||||||||    || ||| | | || ||
Db        578 ATAAGTGAAAAGTTGATTTGGACTTATAAGTTATTAAAAGTGTTTGAATATATATTGA 635


RESULT 5
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match             6.6%;  Score 55.8;  DB 4;  Length 1141;
  Best Local Similarity   8.3%;  Pred. No. 9.8e-05;
  Matches   48;  Conservative  272;  Mismatches  250;  Indels    8;  Gaps    1;

Qy        220 AGCTTTTACATCACTATGGAAGTGATACTATTCGAAGCGGAGTGGAGGAAGAACTTGAGT 279
              :: |:  :: ::: :   : : :: :  : ::: |   ::: : :: :::: |::  :
Db         57 RKWTWARMYCKYRRWYNNKSRWWKGWYKKKWYBCANNTSBRYHARRWKDMKTAYBMTMTN 116

Qy        280 ACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGTAATGTTGAGAG 339
                |:  :  : ::::| :  : :::: |    :  : :  : ::: : :: :     |::
Db        117 KWGKTGWRHRYWRWRAMBDTVDHHYVTAMNNAWTTMCMMDKDDKRTRWWWKKNNNATGWD 176

Qy        340 TATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGG 399
              ::|:::::   | : :|:::: : ::   : :: :  :::: :: :: : ::: ::: ::
Db        177 DDTKYHMWNNNGCBTVTWMVRYKTDRDWSBKRMNYGMBWWKNWSYDVTYYWWVWDDMCKR 236

Qy        400 CAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAA 459
              :::::: :  ::   : ::||:::::     : :|: :  :       |:    : :::
Db        237 KVRRWVRTRGRMRNYMVAWBTAHRRRYNNGWTBAMAYRRWTMNNNNNNAKAMCKRAKYWG 296

Qy        460 AAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAA------- 512
              : :| : : ||| :   |::: : :|   :  :  :::::: ::| : :::
Db        297 WNRABVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDHKWWKWSAAMGVYWNNNNNNNWT 356

Qy        513 -GTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTT 571
               ::|:: | :::::   :::::|   :: :: :    : |::       |:: :::|: ::
Db        357 YKKARHBARWDWVWHSAWKKWHANAAHYSRKKWTBYKRKTMVNNNNGTTMWKRMWAWYWK 416

Qy        572 AGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTA 631
              :::::: |:       : :: :|:  :::    ::      :: ::::|||:  ||:::
Db        417 MDMDWBGTYNNNNNGGRTYYGWTKNKKMWTYYKWKANNCKWRAWDHKTCTHNNTTWWKMK 476

Qy        632 CACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTT 691
                : ::   ::::    :   :::::::       : :: : |: :      :
Db        477 TYWNNCYWKSMTNGKSHRBAAAVYTWYMWWWRRYAHANNNNWDYWWKACTWYKYBVCSKW 536

Qy        692 ATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAA 751
              :  : |: |: ::  |::: ::    :::: :: :  |:|  : : ::: :: |::: :
Db        537 WNNYAAWYTKSSWNYTSRYYRWKTNNSWRWRSDTRSMGRANNYARABHYGYKWNTRWWBW 596

Qy        752 TAGTGTATTTTATTTAAAAAGATCGCATACATTACCAG 789
              : :::     |:::: :: :|:| |::|  :   :||
Db        597 SHTWBHBRAGAAHYWMBMMYBAKCHCMKAWYKAKKYAG 634
```

```
RESULT 6
US-08-998-416-186/c
; Sequence 186, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  186:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  615 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1074RP
US-08-998-416-186

  Query Match             5.6%;  Score 47.4;  DB 3;  Length 615;
  Best Local Similarity   45.4%;  Pred. No. 0.0099;
  Matches  171;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy        456 TTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTA 515
              ||||| || |  |      |   | || |||| || |  | || |       ||  ||   |
Db        587 TTAAATAATGATAAAATATTAATTAAATAATTAATTTAATTAAATAAATAAATAAATAAA 528

Qy        516 ATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCA 575
              || | ||||| ||||||||   | |  |||  ||  | |||   | ||    |      |
Db        527 ATGAATATAATTTATTAAATAATATAAAAATTAATAATAAGAAATTAAAGTTAAAATTAA 468

Qy        576 TTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACA 635
              ||||     |  |||||      |  | || || |     | |  ||| || | |
Db        467 TTTTAATAATAATTCTTATAAAAAGTTAAATAATATAAATCAACATAATATTTATAAAAA 408

Qy        636 AACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATAC 695
              |    | |    ||    ||  |   |       | | |   |  ||   |  | |||
Db        407 TAGATATTATAATAAAAATAATAATTTACAATATTTAAATAAATTAAATCTTTATAATAA 348

Qy        696 TAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGT 755
```

```
              ||||  | |       || | | || | | ||  |||||  | |||  |   |   || |
Db        347 TAAAATATATTATTTTAATAAACAATTAATAAATAATATATTAATTATTGATAATCTATT 288

Qy        756 GTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCACAACTC 815
              ||  | | | |||  |    |||   |   |     | | | | | | ||   || | |
Db        287 TAATAATTTATTAAAGAAAATAATAATATCTAATAATATTTTAATAACTAATTTAAAATT 228

Qy        816 ACAACAAAGTACAAAGA 832
              |||| ||   ||| |
Db        227 TGAACATAGACTAAATA 211


RESULT 7
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match             5.3%;  Score 45.2;  DB 4;  Length 1141;
  Best Local Similarity   10.1%;  Pred. No. 0.044;
  Matches   58;  Conservative  239;  Mismatches  277;  Indels    2;  Gaps    1;

Qy         86 TGGGTCTCAGAAGAGGATCACCTACGAGCAAGTGAAGAAAATGGCAACTTTATTTGATGA 145
              |                    |  ||::    :  :|:   :: ::: :: | |  :  |:::
Db        655 TNNNNNNNNNNNNNNNSCCTCTRMMTMRWTMKGDGMTVRKKVKWRDTTCTYVDVWADSW 596

Qy        146 CTTGTTGATATTTATTGAGAATTACAACTTTGCAGAAAAGCCAACTCTGCGGTTTCAGGT 205
              :::    :   :  :|:     : :|:::::  :  |:: : : |    :: ::||    :
Db        595 VWWYANWMRCRDVTYTRNNTYCKSYAHSYWYWSNNAMWYRRYSARNWSSMARWTTRNNWW 536

Qy        206 TCTGGAATTAATTAAGCTTTTACATCACTATGGAAGTGATACTATTCGAAGCGGAGTGGA 265
              : :::: :| | :::::     :|  :::::::|::  |:::  :      :|  :: :
Db        535 MSGBVRMRWAGTMWWRHWNNNNTDTRYYWWWKRWARBTTTVYDSMCNAKSMWRGNNWRAM 476

Qy        266 GGAAGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGA 325
              : :: |  : || ::  :  : :     ::::|:::: :|:|: |        :| :::: ::
Db        475 KMWWAANNDAGAMDHWTYWMGNNTMWMRRAWKMMNMAWCRRAYCCNNNNNRACVWHKHKM 416

Qy        326 AGTAATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGA 385
              ::: ::  ::|      : |::::: : :  :     :::: |:::  :::| ::|:::|
Db        415 WRWTWKYMWKAACNNNNBKAMYMRVAWMMYSRDTTNTDWMMWTSDWBWHWYTVDYTMMRA 356

Qy        386 CACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGT 445
                  ::  :    ::: : :  :  :|     :| |::  : :: :|   :  :
Db        355 WNNNNNNNWRBCKTTSWMWWMDHMNTHCTYG--NNTWGSAYBMAAMSMWAAGASNBVTYN 298

Qy        446 TTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGAT 505
              : :: |::| |          :   : | | |     :  ::  :: ::: ::  | ::  :
Db        297 WCWRMTYMGKTMTNNNNNNKAWYYRTKTVAWCNNRYYYDTAVWTBKRNYKYCYAYBWYYB 238

Qy        506 AAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGA 565
              :    :: ::: :|::: :  :::: :    : ::  |:: : :|:|:      :::: :
Db        237 MYMGKHHWBWWRRABHRSWNMWWVKCRNKYMVSWHYHAMRYBKWABAVGCNNNWKDRMAH 178

Qy        566 AAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACT 625
```

```
                :::  |   : ::::  :::::   : |:|  : |::: :: :    : :::::   :|:
Db        177 HHWCATNNNMMWWWYAYMHHMHKKGKAAWTNNKTABRDDHBAHVKTYWYWRYDYWCAMCW 118

Qy        626 TTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAA 661
                :  :: |: :   : :: : |  ||::: ::::
Db        117 MNAKAKVRTAMKHMWYYTDRYVSANNTGVRWMMMRW 82


RESULT 8
US-08-960-780-51
; Sequence 51, Application US/08960780
; Patent No. 6204435
;  GENERAL INFORMATION:
;    APPLICANT:  Feitelson, Jerald S.
;    APPLICANT:  Schnepf, H. Ernest
;    APPLICANT:  Narva, Kenneth E.
;    APPLICANT:  Stockhoff, Brian A.
;    APPLICANT:  Schmeits, James
;    APPLICANT:  Loewer, David
;    APPLICANT:  Dullum, Charles Joseph
;    APPLICANT:  Muller-Cohn, Judy
;    APPLICANT:  Stamp, Lisa
;    TITLE OF INVENTION:  No. 6204435el Pesticidal Toxins and Nucleotide
;    TITLE OF INVENTION:  Sequences Which Encode These Toxins
;    NUMBER OF SEQUENCES:  134
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  US
;      ZIP:  32606-6669
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/960,780
;      FILING DATE:  30-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/029,848
;      FILING DATE:  30-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  MA-708
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  352-375-8100
;      TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1341 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  177C8
US-08-960-780-51

  Query Match             5.3%;  Score 44.8;  DB 3;  Length 1341;
  Best Local Similarity   47.7%;  Pred. No. 0.058;
  Matches  163;  Conservative    0;  Mismatches  177;  Indels    2;  Gaps    1;

Qy        269 AGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGT 328
              | | ||| | ||  | |   ||  | |         |   | || |   ||  |||  ||||
Db         67 ATATCTTTATTAAATAATGAAGTGATAAAAGCTGAACAATTAAATATAAATTCTCAAAGT 126

Qy        329 AATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
              ||   |   |  |||||  |   | ||  |   || | |  | || |||||  |  | |
Db        127 AAATATACTAACTTGCAAAATCTAAAAATCACTGACAAGGTAGAGGATTTTAAAGAAGAT 186

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              ||   |||||  | | |  |     |  |    |  ||  | | ||||        | |
Db        187 AAGGAAAAAGCGAAAGAATGGGGGAAAGAAAAAGAAAAAGAGTGGAAACTAACTGCTACT 246

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
              |   |   ||||  | | | || |||  | |   ||   |   |  | ||| |||| |
Db        247 GAAAAAGGAAAAATGAATAATTTTTTAGATAATAAAAATGATATAAAGACAAATTATAAA 306

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              ||| || |  || || |  |        |||| | | |||||| ||| | ||||  |||
Db        307 GAAATTACTTTTTCTATGGCAGGCTCA--TTTGAAGATGAAATAAAAGATTTAAAAGAAA 364

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
              | ||  |  |  |   |  |||  || | |||     | |||
Db        365 TTGATAAGATGTTTGATAAAACCAATCTATCAAATTCTATTA 406


RESULT 9
US-09-073-898-51
; Sequence 51, Application US/09073898
; Patent No. 6242669
;  GENERAL INFORMATION:
;    APPLICANT:  Feitelson, Jerald S.
;    APPLICANT:  Schnepf, H. Ernest
;    APPLICANT:  Narva, Kenneth E.
;    APPLICANT:  Stockhoff, Brian A.
;    APPLICANT:  Schmeits, James
;    APPLICANT:  Loewer, David
;    APPLICANT:  Dullum, Charles Joseph
;    APPLICANT:  Muller-Cohn, Judy
;    APPLICANT:  Stamp, Lisa
;    APPLICANT:  Morrill, George
;    APPLICANT:  Finstad-Lee, Stacey
;    TITLE OF INVENTION:  No. 6242669el Pesticidal Toxins and Nucleotide
;    TITLE OF INVENTION:  Sequences Which Encode These Toxins
;    NUMBER OF SEQUENCES:  144
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  US
;      ZIP:  32606-6669
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/073,898
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/029,848
;      FILING DATE:  30-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/960,780
;      FILING DATE:  30-OCT-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sanders, Jay M.
;      REGISTRATION NUMBER:  39,355
;      REFERENCE/DOCKET NUMBER:  MA-708C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  352-375-8100
;      TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1341 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
```

```
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  PS177C8
US-09-073-898-51

  Query Match             5.3%;  Score 44.8;  DB 3;  Length 1341;
  Best Local Similarity   47.7%;  Pred. No. 0.058;
  Matches  163;  Conservative    0;  Mismatches  177;  Indels    2;  Gaps    1;

Qy        269 AGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGT 328
              | | ||| | ||  | |   ||  | |       |   | || |   ||  |||  |||| 
Db         67 ATATCTTTATTAAATAATGAAGTGATAAAAGCTGAACAATTAAATATAAATTCTCAAAGT 126

Qy        329 AATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
              ||   |   |  ||||| |   | ||  |   || | |  | || |||||  |  | | 
Db        127 AAATATACTAACTTGCAAAATCTAAAAATCACTGACAAGGTAGAGGATTTTAAAGAAGAT 186

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              ||   |||||  | | |  |   |  |     |  ||  | | ||||       | | 
Db        187 AAGGAAAAAGCGAAAGAATGGGGGAAAGAAAAAGAAAAAGAGTGGAAACTAACTGCTACT 246

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
              |   |    ||||  | | | || |||  | |   ||   |     |  | || | |||| 
Db        247 GAAAAAGGAAAAATGAATAATTTTTTAGATAATAAAAATGATATAAAGACAAATTATAAA 306

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              ||| || |  || || | |         |||| | | |||||| ||| | ||||  ||| 
Db        307 GAAATTACTTTTTCTATGGCAGGCTCA--TTTGAAGATGAAATAAAAGATTTAAAAGAAA 364

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
              | | |  |  |   |  |||  || | |||     | |||
Db        365 TTGATAAGATGTTTGATAAAACCAATCTATCAAATTCTATTA 406


RESULT 10
US-09-371-913A-2
; Sequence 2, Application US/09371913A
; Patent No. 6297369
; GENERAL INFORMATION:
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Narva, Kenneth E.
;  APPLICANT:  Stockhoff, Brian A.
;  APPLICANT:  Finstad Lee, Stacey
;  APPLICANT:  Walz, Mikki
;  APPLICANT:  Sturgis, Blake
;  TITLE OF INVENTION:  Pesticidal Toxins and Genes from Bacillus laterosporus
;  TITLE OF INVENTION:  Strains
;  FILE REFERENCE:  MA-719XC2 US
;  CURRENT APPLICATION NUMBER: US/09/371,913A
;  CURRENT FILING DATE:  1999-08-10
;  PRIOR APPLICATION NUMBER:  60/095,955
;  PRIOR FILING DATE:  1998-08-10
;  PRIOR APPLICATION NUMBER:  60/138,251
;  PRIOR FILING DATE:  1999-06-08
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1341
;   TYPE: DNA
;   ORGANISM: Bacillus laterosporus
US-09-371-913A-2

  Query Match             5.3%;  Score 44.8;  DB 3;  Length 1341;
  Best Local Similarity   47.7%;  Pred. No. 0.058;
  Matches  163;  Conservative    0;  Mismatches  177;  Indels    2;  Gaps    1;

Qy        269 AGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGT 328
              | | ||| | ||  | |   ||  | |       |   | || |   ||  |||  |||| 
Db         67 ATATCTTTATTAAATAATGAAGTGATAAAAGCTGAACAATTAAATATAAATTCTCAAAGT 126

Qy        329 AATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
              ||   |   |  |||||  |   | ||  |   || | |  | || |||||  |  | | 
Db        127 AAATATACTAACTTGCAAAATCTAAAAATCACTGACAAGGTAGAGGATTTTAAAGAAGAT 186

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              ||   |||||  | | |  |   |  |     |  ||  | | ||||       | | 
Db        187 AAGGAAAAAGCGAAAGAATGGGGGAAAGAAAAAGAAAAAGAGTGGAAACTAACTGCTACT 246

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
              |   |    ||||  | | | || |||  | |   ||   |     |  | || | |||| 
Db        247 GAAAAAGGAAAAATGAATAATTTTTTAGATAATAAAAATGATATAAAGACAAATTATAAA 306

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              ||| || |  || || | |         |||| | | |||||| ||| | ||||  ||| 
Db        307 GAAATTACTTTTTCTATGGCAGGCTCA--TTTGAAGATGAAATAAAAGATTTAAAAGAAA 364

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
              | | |  |  |   |  |||  || | |||     | |||
Db        365 TTGATAAGATGTTTGATAAAACCAATCTATCAAATTCTATTA 406


RESULT 11
US-09-307-106-11
; Sequence 11, Application US/09307106
; Patent No. 6603063
; GENERAL INFORMATION:
;  APPLICANT:  Feitelson, Jerald S.
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Narva, Kenneth E.
;  APPLICANT:  Stockhoff, Brian A.
;  APPLICANT:  Schmeits, James
;  APPLICANT:  Loewer, David
;  APPLICANT:  Dullum, Charles Joseph
;  APPLICANT:  Muller-Cohn, Judy
;  APPLICANT:  Stamp, Lisa
;  APPLICANT:  Morrill, George
;  APPLICANT:  Finstad-Lee, Stacey
;  TITLE OF INVENTION:  No. 6603063el Pesticidal Toxins and Nucleotide
;  TITLE OF INVENTION:  Sequences Which Encode These Toxins
;  NUMBER OF SEQUENCES:  54
;  CORRESPONDENCE ADDRESS:
;    ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;    STREET:  2421 N.W. 41st Street, Suite A-1
;    CITY:  Gainesville
;    STATE:  FL
;    COUNTRY:  US
;    ZIP:  32606-6669
;  COMPUTER READABLE FORM:
;    MEDIUM TYPE:  Floppy disk
;    COMPUTER:  IBM PC compatible
;    OPERATING SYSTEM:  PC-DOS/MS-DOS
;    SOFTWARE:  PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/09/307,106
;    FILING DATE:
;    CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 60/029,848
;    FILING DATE:  30-OCT-1996
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 08/960,780
;    FILING DATE:  30-OCT-1997
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 09/073,898
;    FILING DATE:  05-MAY-1998
;  ATTORNEY/AGENT INFORMATION:
;    NAME:  Sanders, Jay M.
;    REGISTRATION NUMBER:  39,355
;    REFERENCE/DOCKET NUMBER:  MA-708C2
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  352-375-8100
;    TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1341 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  PS177C8a
US-09-307-106-11

  Query Match             5.3%;  Score 44.8;  DB 4;  Length 1341;
  Best Local Similarity   47.7%;  Pred. No. 0.058;
  Matches  163;  Conservative    0;  Mismatches  177;  Indels    2;  Gaps    1;

Qy        269 AGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGT 328
              | | ||| | ||  | |   ||  | |        |   | || |   ||  |||  ||||
Db         67 ATATCTTTATTAAATAATGAAGTGATAAAAGCTGAACAATTAAATATAAATTCTCAAAGT 126

Qy        329 AATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
              ||   |   |  |||||  |  | ||  |  || | |  | || |||||  | | |
Db        127 AAATATACTAACTTGCAAAATCTAAAAATCACTGACAAGGTAGAGGATTTTAAAGAAGAT 186

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              ||   |||||  | | |  |   |  |  ||  | | ||||       | |
Db        187 AAGGAAAAAGCGAAAGAATGGGGGAAAGAAAAAGAAAAAGAGTGGAAACTAACTGCTACT 246

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
              |   |    ||||  | | | || |||  | |   ||   |    |  | || | ||||
Db        247 GAAAAAGGAAAAATGAATAATTTTTTAGATAATAAAAATGATATAAAGACAAATTATAAA 306

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              |||  || |  || || | |          |||| | | ||||||  ||| | ||||  |||
Db        307 GAAATTACTTTTTCTATGGCAGGCTCA--TTTGAAGATGAAATAAAAGATTTAAAAGAAA 364

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
              | |  |  |  |   |  |||  || | |||      | |||
Db        365 TTGATAAGATGTTTGATAAAACCAATCTATCAAATTCTATTA 406


RESULT 12
US-09-967-805-2
; Sequence 2, Application US/09967805
; Patent No. 6605701
; GENERAL INFORMATION:
;  APPLICANT: Schnepf, Ernest H
;  APPLICANT:  Narva, Kenneth E
;  APPLICANT:  Stockhoff, Brian A
;  APPLICANT:  Finstad Lee, Stacey
;  APPLICANT:  Walz, Mikki
;  APPLICANT:  Sturgis, Blake
;  TITLE OF INVENTION: Pesticidal Toxins and Genes from Bacillus laterosporus
;  TITLE OF INVENTION:  Strains
;  FILE REFERENCE: MA-719XC2D1
;  CURRENT APPLICATION NUMBER: US/09/967,805
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: 09/371,913
;  PRIOR FILING DATE: 1999-08-10
;  PRIOR APPLICATION NUMBER: 60/095,955
;  PRIOR FILING DATE: 1998-08-10
;  PRIOR APPLICATION NUMBER: 60/138,251
;  PRIOR FILING DATE: 1999-06-08
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1341
;   TYPE: DNA
;   ORGANISM: Bacillus laterosporus
US-09-967-805-2

  Query Match             5.3%;  Score 44.8;  DB 4;  Length 1341;
  Best Local Similarity   47.7%;  Pred. No. 0.058;
  Matches  163;  Conservative    0;  Mismatches  177;  Indels    2;  Gaps    1;

Qy        269 AGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGT 328
              | | ||| | ||  | |   ||  | |        |   | || |   ||  |||  ||||
Db         67 ATATCTTTATTAAATAATGAAGTGATAAAAGCTGAACAATTAAATATAAATTCTCAAAGT 126

Qy        329 AATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
              ||   |   |  |||||  |  | ||  |  || | |  | || |||||  | | |
Db        127 AAATATACTAACTTGCAAAATCTAAAAATCACTGACAAGGTAGAGGATTTTAAAGAAGAT 186

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              ||   |||||  | | |  |   |  |  ||  | | ||||       | |
Db        187 AAGGAAAAAGCGAAAGAATGGGGGAAAGAAAAAGAAAAAGAGTGGAAACTAACTGCTACT 246

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
              |   |    ||||  | | | || |||  | |   ||   |    |  | || | ||||
Db        247 GAAAAAGGAAAAATGAATAATTTTTTAGATAATAAAAATGATATAAAGACAAATTATAAA 306

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              |||  || |  || || | |          |||| | | ||||||  ||| | ||||  |||
Db        307 GAAATTACTTTTTCTATGGCAGGCTCA--TTTGAAGATGAAATAAAAGATTTAAAAGAAA 364

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
              | |  |  |  |   |  |||  || | |||      | |||
Db        365 TTGATAAGATGTTTGATAAAACCAATCTATCAAATTCTATTA 406


RESULT 13
US-09-850-351A-51
; Sequence 51, Application US/09850351A
; Patent No. 6656908
;  GENERAL INFORMATION:
;       APPLICANT: Feitelson, Jerald S.
;                  Schnepf, H. Ernest
;                  Narva, Kenneth E.
;                  Stockhoff, Brian A.
;                  Schmeits, James
;                  Loewer, David
;                  Dullum, Charles Joseph
;                  Muller-Cohn, Judy
;                  Stamp, Lisa
;                  Morrill, George
;       TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide
;                           Sequences Which Encode These Toxins
;       NUMBER OF SEQUENCES: 144
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
;            STREET: 2421 N.W. 41st Street, Suite A-1
;            CITY: Gainesville
;            STATE: FL
;            COUNTRY: US
;            ZIP: 32606-6669
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/850,351A
;            FILING DATE: 07-May-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 09/073,898
;            FILING DATE: 06-MAY-1998
;            APPLICATION NUMBER: US 08/960,780
;            FILING DATE: 30-OCT-1997
;            APPLICATION NUMBER: US 60/029,848
;            FILING DATE: 30-OCT-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Sanders, Jay M.
;            REGISTRATION NUMBER: 39,355
;            REFERENCE/DOCKET NUMBER: MA-708CD1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 352-375-8100
;            TELEFAX: 352-372-5800
;  INFORMATION FOR SEQ ID NO: 51:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1341 base pairs
;            TYPE: nucleic acid
```

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        ORIGINAL SOURCE:
;             INDIVIDUAL ISOLATE: PS177C8
;        SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-850-351A-51

  Query Match             5.3%;  Score 44.8;  DB 4;  Length 1341;
  Best Local Similarity   47.7%;  Pred. No. 0.058;
  Matches  163;  Conservative    0;  Mismatches  177;  Indels    2;  Gaps    1;

Qy        269 AGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGT 328
              | | ||| | ||  | |   ||  | |       |   | || |   ||  |||  |||| 
Db         67 ATATCTTTATTAAATAATGAAGTGATAAAAGCTGAACAATTAAATATAAATTCTCAAAGT 126

Qy        329 AATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
              ||   |   |   ||||| |   | ||  |   || | |  | || ||||| |  | | 
Db        127 AAATATACTAACTTGCAAAATCTAAAAATCACTGACAAGGTAGAGGATTTTAAAGAAGAT 186

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              ||   |||||  | | |  |    | |    |  ||  | | ||||       | | 
Db        187 AAGGAAAAAGCGAAAGAATGGGGGAAAGAAAAAGAAAAAGAGTGGAAACTAACTGCTACT 246

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
              |    |   |||| | | | || ||| | |   ||   |     |  | || | ||||
Db        247 GAAAAAGGAAAAATGAATAATTTTTTAGATAATAAAAATGATATAAAGACAAATTATAAA 306

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              ||| || |  || || |  |       |||| | | |||||| ||| | ||||  ||| 
Db        307 GAAATTACTTTTTCTATGGCAGGCTCA--TTTGAAGATGAAATAAAAGATTTAAAAGAAA 364

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
              | |  |  |  |   |  |||  || | |||     | |||
Db        365 TTGATAAGATGTTTGATAAAACCAATCTATCAAATTCTATTA 406


RESULT 14
US-08-471-033-22
; Sequence 22, Application US/08471033
; Patent No. 5770696
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  No. 5770696el Pesticidal Proteins and Strains
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CIBA-GEIGY Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,033
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  P-40,403
;      REFERENCE/DOCKET NUMBER:  CGC 1695/CIP3/DIV7 - SQLv3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8582
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4041 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4038
;      OTHER INFORMATION:  /product= "VIP1A(a)/VIP2A(a) fusion
;      OTHER INFORMATION:  product"
US-08-471-033-22

  Query Match             5.3%;  Score 44.8;  DB 1;  Length 4041;
  Best Local Similarity   47.7%;  Pred. No. 0.086;
  Matches  163;  Conservative    0;  Mismatches  177;  Indels    2;  Gaps    1;

Qy        269 AGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGT 328
              | | ||| | ||  | |   ||  | |       |   | || |   ||  |||  |||| 
Db         88 ATATCTTTATTAAATAATGAAGTGATAAAAGCTGAACAATTAAATATAAATTCTCAAAGT 147

Qy        329 AATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
              ||   |   |   |||| |   | ||  |   || | |  | || ||||| |  | | 
Db        148 AAATATACTAACTTGCAAAATCTAAAAATCACTGACAAGGTAGAGGATTTTAAAGAAGAT 207

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              ||   |||||  | | |  |    | |    |  ||  | | ||||       | | 
Db        208 AAGGAAAAAGCGAAAGAATGGGGGAAAGAAAAAGAAAAAGAGTGGAAACTAACTGCTACT 267

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
              |    |   |||| | | | || ||| | |   ||   |     |  | || | ||||
Db        268 GAAAAAGGAAAAATGAATAATTTTTTAGATAATAAAAATGATATAAAGACAAATTATAAA 327

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              ||| || |  || || |  |       |||| | | |||||| ||| | ||||  ||| 
Db        328 GAAATTACTTTTTCTATGGCAGGCTCA--TTTGAAGATGAAATAAAAGATTTAAAAGAAA 385

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
              | |  |  |  |   |  |||  || | |||     | |||
Db        386 TTGATAAGATGTTTGATAAAACCAATCTATCAAATTCTATTA 427


RESULT 15
US-08-471-044-22
; Sequence 22, Application US/08471044
; Patent No. 5840868
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  No. 5840868el Pesticidal Proteins and Strains
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CIBA-GEIGY Corporation
```

```
;       STREET:  7 Skyline Drive
;       CITY:  Hawthorne
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10532
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/471,044
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/463,483
;       FILING DATE:  05-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/314,594
;       FILING DATE:  09-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/218,018
;       FILING DATE:  23-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/037,057
;       FILING DATE:  25-MAR-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Pace, Gary M.
;       REGISTRATION NUMBER:  40,403
;       REFERENCE/DOCKET NUMBER:  CGC 1695/CIP3/DIV6 - SQLv3
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  919-541-8582
;       TELEFAX:  919-541-8689
;   INFORMATION FOR SEQ ID NO:  22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4041 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..4038
;       OTHER INFORMATION:  /product= "VIP1A(a)/VIP2A(a) fusion
;       OTHER INFORMATION:  product"
US-08-471-044-22

  Query Match             5.3%;  Score 44.8;  DB 2;  Length 4041;
  Best Local Similarity   47.7%;  Pred. No. 0.086;
  Matches  163;  Conservative    0;  Mismatches  177;  Indels    2;  Gaps    1;

Qy        269 AGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGT 328
              | | ||| | ||  | |   ||  | |        |   | || |   ||  |||  ||||
Db         88 ATATCTTTATTAAATAATGAAGTGATAAAAGCTGAACAATTAAATATAAATTCTCAAAGT 147

Qy        329 AATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
              ||   |   |  |||||  |   | ||  |   || | |  | || |||||  |  | |
Db        148 AAATATACTAACTTGCAAAATCTAAAAATCACTGACAAGGTAGAGGATTTTAAAGAAGAT 207

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              ||   |||||  | | |  |    |  |    |  ||  | | ||||       | |
Db        208 AAGGAAAAAGCGAAAGAATGGGGGAAAGAAAAAGAAAAAGAGTGGAAACTAACTGCTACT 267

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
              |   |    ||||  | | | || |||  | |   ||   |    |  | || | ||||
Db        268 GAAAAAGGAAAAATGAATAATTTTTTAGATAATAAAAATGATATAAAGACAAATTATAAA 327

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
               ||| || |  || || |  |         |||| | | |||||| ||| | ||||  |||
Db        328 GAAATTACTTTTTCTATGGCAGGCTCA--TTTGAAGATGAAATAAAAGATTTAAAAGAAA 385

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
               | |  |  |  |   |  |||  || | |||     | |||

Db        386 TTGATAAGATGTTTGATAAAACCAATCTATCAAATTCTATTA 427


Search completed: December  6, 2004, 22:52:27
Job time : 86.3031 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 19:02:39 ; Search time 502.824 Seconds
(without alignments)
9300.077 Million cell updates/sec

Title: US-09-806-197-2
Perfect score: 851
Sequence: 1 ctgaaaaggaagttcatcga.....aaaaagttgcaattctgtca 851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3695051 seqs, 2747533894 residues

Total number of hits satisfying chosen parameters: 7390102

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	55.2	6.5	3673778	15	US-10-312-841-1	Sequence 1, Appli
c 2	54.6	6.4	2114	16	US-10-257-166-162	Sequence 162, App
3	52.8	6.2	4045	16	US-10-240-589C-130	Sequence 130, App
4	52.2	6.1	8056	18	US-10-473-126-386	Sequence 386, App
5	52.2	6.1	3673778	15	US-10-312-841-2	Sequence 2, Appli
c 6	52	6.1	539	10	US-09-814-353-5852	Sequence 5852, Ap
c 7	52	6.1	539	10	US-09-814-353-12133	Sequence 12133, A
c 8	52	6.1	13712	15	US-10-311-455-1504	Sequence 1504, Ap
c 9	51.8	6.1	19787	15	US-10-311-455-1424	Sequence 1424, Ap
c 10	51.6	6.1	8676	16	US-10-221-613-378	Sequence 378, App
c 11	50.6	5.9	802	18	US-10-425-115-647	Sequence 647, App
c 12	49.8	5.9	9157	15	US-10-311-455-1473	Sequence 1473, Ap
c 13	49.6	5.8	17721	15	US-10-311-455-1701	Sequence 1701, Ap
c 14	49.4	5.8	578	18	US-10-425-115-160787	Sequence 160787,
c 15	49.2	5.8	113515	15	US-10-311-455-2148	Sequence 2148, Ap
c 16	48.6	5.7	628	16	US-10-424-599-75713	Sequence 75713, A
c 17	48.6	5.7	3673778	15	US-10-312-841-2	Sequence 2, Appli
c 18	48.4	5.7	8404	15	US-10-311-455-1568	Sequence 1568, Ap
c 19	48.4	5.7	8404	16	US-10-221-714A-222	Sequence 222, App
c 20	48	5.6	6378	15	US-10-311-455-149	Sequence 149, App
c 21	48	5.6	6378	17	US-10-433-793-57	Sequence 57, Appl
22	47.8	5.6	14924	15	US-10-311-455-198	Sequence 198, App
23	47.8	5.6	14924	15	US-10-240-452-22	Sequence 22, Appl
c 24	47.8	5.6	19380	16	US-10-221-613-390	Sequence 390, App
c 25	47.4	5.6	7195	14	US-10-239-676-30	Sequence 30, Appl
c 26	47.4	5.6	7195	15	US-10-240-453-40	Sequence 40, Appl
27	47.4	5.6	9180	15	US-10-311-455-1938	Sequence 1938, Ap
c 28	46.8	5.5	5689	14	US-10-239-676-90	Sequence 90, Appl
c 29	46.8	5.5	5689	15	US-10-240-453-100	Sequence 100, App
c 30	46.8	5.5	5689	16	US-10-221-714A-148	Sequence 148, App
31	46.6	5.5	40862	15	US-10-311-455-2046	Sequence 2046, Ap
32	46.4	5.5	5145	15	US-10-311-455-322	Sequence 322, App
33	46.4	5.5	5145	15	US-10-240-485-18	Sequence 18, Appl
c 34	46.4	5.5	6255	15	US-10-311-455-2001	Sequence 2001, Ap
c 35	46.4	5.5	21537	15	US-10-311-455-1971	Sequence 1971, Ap
36	46.4	5.5	3673778	15	US-10-312-841-1	Sequence 1, Appli
37	46	5.4	6069	16	US-10-221-613-343	Sequence 343, App
38	46	5.4	16217	15	US-10-311-455-597	Sequence 597, App
39	45.8	5.4	8056	18	US-10-473-126-240	Sequence 240, App
40	45.8	5.4	38342	16	US-10-221-714A-472	Sequence 472, App
c 41	45.8	5.4	513509	10	US-09-754-853A-4	Sequence 4, Appli
c 42	45.4	5.3	363	18	US-10-425-115-83210	Sequence 83210, A
43	45.2	5.3	586	16	US-10-424-599-80168	Sequence 80168, A
c 44	45.2	5.3	8056	18	US-10-473-126-386	Sequence 386, App
c 45	45	5.3	5989	16	US-10-221-714A-238	Sequence 238, App

ALIGNMENTS

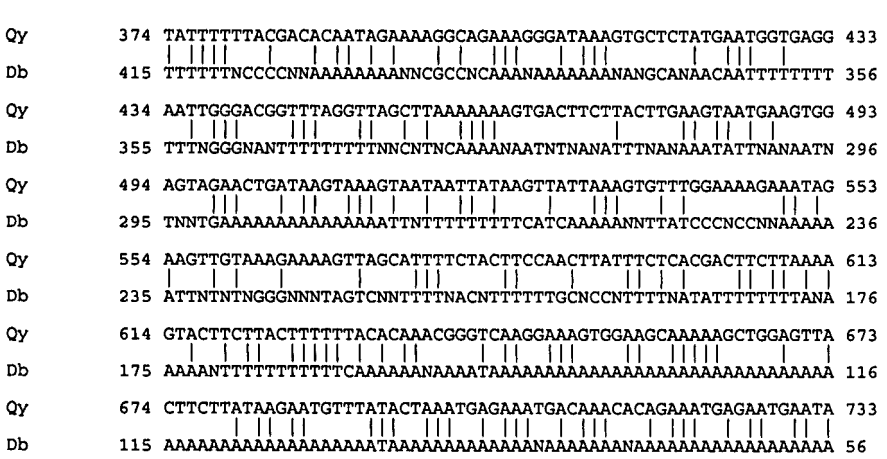
RESULT 1
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MH
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 6.5%; Score 55.2; DB 15; Length 3673778;
Best Local Similarity 48.6%; Pred. No. 1.1;
Matches 179; Conservative 0; Mismatches 188; Indels 1; Gaps 1;

```
Qy      456 TTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTA 515
            | ||||| |   |     | |  ||  ||  ||   | ||  | |  |||    |   |
Db    84258 TCAAAAATAAAAAAAAAAAATTAAAAACAACAAATATAACTAATATTATTAAAATACAAA 84199

Qy      516 ATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCA 575
            | | | | || |   ||||   | |  | || ||||| ||  |  ||| ||   | |  |
Db    84198 AAATTAAAAAATATATAAATAATATCCATAAAAAATAAAAACTTAAAAAAACCTTAATAA 84139

Qy      576 TTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACA 635
```

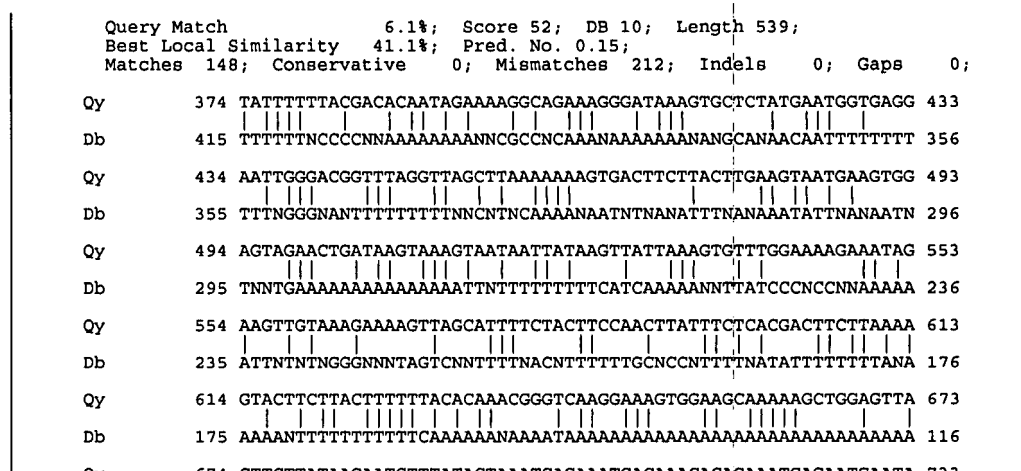
```
Db      84138 TTATCTAATACTATATCACAAACTACCTCAAAATTTAATAACTTAATAAAACAAATATCT 84079

Qy        636 AACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATAC 695

Db      84078 ATTAACTCAAAAAAATACTACACACACAAATAAAATAAATTC-TATAAAAATAACAAAAA 84020

Qy        696 TAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGT 755

Db      84019 TTAAATAAAAAAAAAAATTTTACAAATTAAAATAAAAATTATTAATATACAAATAAAAAT 83960

Qy        756 GTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCACAACTC 815

Db      83959 AATTTTTAATCCAAATCCTTCAAAAAACAAATACCAAAATAAACTTAAAATCTACGACAT 83900

Qy        816 ACAACAAA 823

Db      83899 CTTATAAA 83892


RESULT 2
US-10-257-166-162/c
; Sequence 162, Application US/10257166
; Publication No. US20040023230A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
;  TITLE OF INVENTION:  Genes Implicated in Pharmacogenomics
;  FILE REFERENCE: 5013.1011
;  CURRENT APPLICATION NUMBER: US/10/257,166
;  CURRENT FILING DATE:  2002-10-07
;  PRIOR APPLICATION NUMBER: PCT/EP01/07470
;    DE 10032529.7
;    DE 10043826.1
;  PRIOR FILING DATE: 2001-06-29
;    2000-06-30
;    2000-09-01
;  NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 162
;   LENGTH: 2114
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-162

  Query Match             6.4%;  Score 54.6;  DB 16;  Length 2114;
  Best Local Similarity   56.4%;  Pred. No. 0.068;
  Matches  102;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy        649 AAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATG 708

Db       1499 AAATAAAAACCAAAAAATATATTTAATAATCATAATTATATATATATAAAATAAAAACA 1440

Qy        709 ACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAA 768

Db       1439 ACAATAATACAAAAAATAAATAAATAAAATTAAAAATATCTTATTATTTATTTATTTTAA 1380

Qy        769 AAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACA 828

Db       1379 AATAAAATCTTACTTTACTACCCAAATTAAAATACAATAAAATAATCTCAACTCACTACA 1320

Qy        829 A 829

Db       1319 A 1319


RESULT 3
US-10-240-589C-130
; Sequence 130, Application US/10240589C
; Publication No. US20040076956A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with
;  TITLE OF INVENTION:  DNA repair
;  FILE REFERENCE: 5013.1008
;  CURRENT APPLICATION NUMBER: US/10/240,589C
;  CURRENT FILING DATE:  2003-09-02
;  PRIOR APPLICATION NUMBER: PCT/EP01/03972
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 148
; SEQ ID NO 130
;   LENGTH: 4045
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-589C-130

  Query Match             6.2%;  Score 52.8;  DB 16;  Length 4045;
  Best Local Similarity   47.5%;  Pred. No. 0.22;
  Matches  188;  Conservative    0;  Mismatches  207;  Indels    1;  Gaps    1;

Qy        374 TATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGG 433

Db        858 TTTGTTTTAAAAGAAAATATATTTTGTAAAAAAAAAAAAAGAAAAAAAAGAAAAGAAAAA 917

Qy        434 AATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGG 493

Db        918 AAGGGGGGGAGTGTAGTGGAAATTGAGTATAATAAAATGTTAAGTGTTTTAATAAAATAA 977

Qy        494 AGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAG 553

Db        978 AATAAACGTAAAATGTTTTGAGGGGAGAGTTTTTTTTAAAGTATTTGTTAAAGTTAGAG 1037

Qy        554 AAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAA 613

Db       1038 AAGAAATAATGGGTATTTTGTTAAGTTAGGGTAAGAATATTTTTTAAAATTTTATAAGTG 1097

Qy        614 GTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAG-CAAAAAGCTGGAGTT 672

Db       1098 TAAATTTTTTTAGTTTTAAATATTTATATAAGTTTTATTGGAAGATAAATATATTAGAGT 1157

Qy        673 ACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAAT 732

Db       1158 AATTTTAATATGGTTTAAATTATGTAATGTTAAATAATTTTTTTAAAAAAAGATTTTAAG 1217

Qy        733 ATGATTATTGGTTTAATAATAGTGTATTTTATTTAA 768

Db       1218 TAAAATGTATGGTTTATGAATAGAAATTATGTTTTA 1253


RESULT 4
US-10-473-126-386
; Sequence 386, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
;  APPLICANT: Epigenomics AG
;  TITLE OF INVENTION: Methods and nucleic acids for the analysis of  hematopoietic cell
;  TITLE OF INVENTION:  proliferative disorders
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/473,126
;  CURRENT FILING DATE:  2003-09-26
;  NUMBER OF SEQ ID NOS: 1258
```

```
; SEQ ID NO 386
;   LENGTH: 8056
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-386

  Query Match             6.1%;  Score 52.2;  DB 18;  Length 8056;
  Best Local Similarity   44.8%;  Pred. No. 0.4;
  Matches  201;  Conservative    0;  Mismatches  248;  Indels    0;  Gaps    0;

Qy        375 ATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGA 434
              ||||| ||       ||| ||||   |  ||  | || | |  | |||||   |||
Db       3023 ATTTTATAAATATAAATTAAAAATAAAATAAATGTTATAATTGTGAATGAAATTTGAATT 3082

Qy        435 ATGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGA 494
              ||||  |   ||  |  ||  |||| ||||      | ||  |  ||  |||  | ||
Db       3083 ATTGTTATTTTTATGTGTAATTTAATAAAAAATTAATTTTTTTAAAATGAATTTAATGAT 3142

Qy        495 GTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGA 554
              |    |  ||   |  |  |||||| ||| | || | | | ||   ||||    |
Db       3143 GATTGAAAAATTGTTTTATATATAATTTTAAATAATAATATTTTTAATAAAAATTTTGTT 3202

Qy        555 AGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAG 614
              ||| |  | |||   |   ||   |  ||  |  || |||         |  |||  |
Db       3203 GTTTGAATTGTAAATAAAATTTTAATTTATTATATTTTTTTTTAATAATTTGATTATTAT 3262

Qy        615 TACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTAC 674
              |  | |||  ||  ||  | |||     |      |  || |  |  |     | ||   |
Db       3263 TTTTAATTAATTGATTGTAAAAAAATTTTTGTTATTGTTGTTGTTATTTTTTTGAAAAAA 3322

Qy        675 TTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATAT 734
                 |||  | |||      |||   |||   |    | |  |     ||  ||||
Db       3323 AAAAAATATAATTGTAATGTAAAAAAATGAATAAAATTTGAAATGATGATTAAAAAATAA 3382

Qy        735 GATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGA 794
              | |          |  |   || ||||| ||||||  |   |    || ||    |     |
Db       3383 AAATGAAATAAATAATAATTTGAATTTTTTTTAAATATTTTTGTTAAATGTTGAAAATTA 3442

Qy        795 TGAAGTTATTCATCACAACTCACAACAAA 823
              | |  |  || || |  | |   || |||
Db       3443 TAATTTGTTTTATTATTATTATTAATAAA 3471


RESULT 5
US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
;  APPLICANT: Epigenomics AG
;  TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
;  FILE REFERENCE: E01/1208/WO
;  CURRENT APPLICATION NUMBER: US/10/312,841
;  CURRENT FILING DATE:  2002-12-30
;  NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
;   LENGTH: 3673778
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (379615)
US-10-312-841-2

  Query Match             6.1%;  Score 52.2;  DB 15;  Length 3673778;
  Best Local Similarity   48.2%;  Pred. No. 5.2;
  Matches  147;  Conservative    0;  Mismatches  158;  Indels    0;  Gaps    0;

Qy        470 TTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTA 529
              ||  |  |  ||   |||||||  || |    |   ||||||  | |   | |||   |

Db    3322298 TTTGTTTTAAAAAATATGAAGTTAAGAATTTATTTAAAGTAATTTTAGTTTGATAGTGTT 3322357

Qy        530 TTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAA 589
              |||| || ||| ||||||||||| || |  | |  | |   ||    |||    ||  ||
Db    3322358 TTAAGGTATTTAGAAAAGAAATATAAATGTTGAGAATATAGTATTCGTTTAAGGTTTTAA 3322417

Qy        590 CTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAA 649
                |||| |   |  |  |   |  |  |  | ||| |     |||    | ||  ||
Db    3322418 AGAATTTTTATAGTTTAAGTTTTAAAAAATTCGAATTTATAGTTAAAAATTATAAAATAA 3322477

Qy        650 AGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGA 709
              ||   ||    |      |||  |      | |  || | |  || ||   || | || |
Db    3322478 AGAAAAAAAGTATTATGAGAGAGATAGTAGAAATAAAAGATAGTAGTATTAGATATATAA 3322537

Qy        710 CAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAA 769
              |  |  ||| || ||||| |  |    ||     | |  | ||   |   | ||||||
Db    3322538 AATTCGTAGATATTAGAATTATGAGATATAGAATATAATAATTATATTTAATATTTTAAA 3322597

Qy        770 AAGAT 774
              | ||
Db    3322598 TATAT 3322602


RESULT 6
US-09-814-353-5852/c
; Sequence 5852, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
;  APPLICANT: Lee, John
;  APPLICANT:  Thompson, Pamela
;  APPLICANT:  Lillie, James
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF OVARIAN CANCER
;  FILE REFERENCE: MRI-006B
;  CURRENT APPLICATION NUMBER: US/09/814,353
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: US 60/191,031
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: US 60/207,124
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 60/211,940
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: US 60/216,820
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: US 60/220,661
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: US 60/257,672
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 22037
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5852
;   LENGTH: 539
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 74, 82, 152, 171, 177, 190, 195, 198, 207, 210, 215, 216,
;   LOCATION: 222, 223, 224, 228, 230, 232, 241, 242, 245, 253, 254, 273,
;   LOCATION: 293, 294, 296, 300, 302, 310, 312, 317, 319, 321, 325, 331,
;   LOCATION: 333, 335, 336, 346, 348, 352, 369, 373, 375, 383, 388
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 393, 394, 403, 404, 409, 423, 425, 428, 443, 446, 447, 449,
;   LOCATION: 451, 455, 465, 468, 469, 472, 474, 475, 478, 481, 488, 490,
;   LOCATION: 491, 495, 496, 504, 513, 515, 518, 525, 526
;   OTHER INFORMATION: n = A,T,C or G
US-09-814-353-5852

  Query Match             6.1%;  Score 52;  DB 10;  Length 539;
  Best Local Similarity   41.1%;  Pred. No. 0.15;
  Matches  148;  Conservative    0;  Mismatches  212;  Indels    0;  Gaps    0;
```



RESULT 7
US-09-814-353-12133/c
; Sequence 12133, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12133
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 74, 82, 152, 171, 177, 190, 195, 198, 207, 210, 215, 216,
; LOCATION: 222, 223, 224, 228, 230, 232, 241, 242, 245, 253, 254, 273,
; LOCATION: 293, 294, 296, 300, 302, 310, 312, 317, 319, 321, 325, 331,
; LOCATION: 333, 335, 336, 346, 348, 352, 369, 373, 375, 383, 388
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 393, 394, 403, 404, 409, 423, 425, 428, 443, 446, 447, 449,
; LOCATION: 451, 455, 465, 468, 469, 472, 474, 475, 478, 481, 488, 490,
; LOCATION: 491, 495, 496, 504, 513, 515, 518, 525, 526
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-12133

Query Match 6.1%; Score 52; DB 10; Length 539;
Best Local Similarity 41.1%; Pred. No. 0.15;
Matches 148; Conservative 0; Mismatches 212; Indels 0; Gaps 0;



RESULT 8
US-10-311-455-1504/c
; Sequence 1504, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1504
; LENGTH: 13712
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1504

Query Match 6.1%; Score 52; DB 15; Length 13712;
Best Local Similarity 50.0%; Pred. No. 0.56;
Matches 157; Conservative 0; Mismatches 155; Indels 2; Gaps 1;

```
Qy        696 TAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGT 755
              ||   || |   ||||| | | |  | ||| || || |  |||  |||  |  |  |
Db       9911 CTCATTTAAATTATAAAACAAAAATAAAAAAATTAAAATTA--ATTCTTTTTTTTTTTT 9854

Qy        756 GTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCACAACTC 815
              | |||| |||||||  |   |  ||  || || ||| |  ||  ||          | ||
Db       9853 TTTTTTTTTTTAAAACAAAATCTCACTCTATCACCCAAACTAAAATACAATAATATAATC 9794

Qy        816 ACAACAAAGTACAA 829
              |||| || |||||
Db       9793 TCAACTAACTACAA 9780


RESULT 9
US-10-311-455-1424/c
; Sequence 1424, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1424
;   LENGTH: 19787
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1424

  Query Match             6.1%;  Score 51.8;  DB 15;  Length 19787;
  Best Local Similarity   43.6%;  Pred. No. 0.72;
  Matches  230;  Conservative    0;  Mismatches  297;  Indels    0;  Gaps    0;

Qy        310 ACAAGAAAGCTCTGGAAGTAATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTC 369
              | ||| |   |||  || | | |        ||   | ||  | ||||     |||  |
Db       3574 AAAATATCTTTCTTTAATTTTTATCACCTTCATTTCAAAATAAAAATAATAATCTACTAC 3515

Qy        370 AAAGTATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGT 429
              ||| | ||| |      || |||  |||       |      ||  | | |   |  |||
Db       3514 AAAATTTTTATATTATCATAATTCAAATAAACAATTTCAATTCATTTCATAAAAAATAAC 3455

Qy        430 GAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAA 489
              |  ||  |     ||   |||  || |||| |   | | ||        | |  |
Db       3454 TAAAAAACGATTATATTCTTTTTATTTTTAAAATATAAAAATTTTTAAAATTCTTAATAT 3395

Qy        490 GTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAA 549
              | |  ||      |    || ||  ||   | || ||   ||   |  | ||  | ||
Db       3394 AATAACTTAAAAAATAATAATAAATACAAAAATTTCATAATATAATCCTTAATATCAAAA 3335

Qy        550 ATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTT 609
              |||  |   || | |||  |   | |   | ||  |||  |      |    |
Db       3334 ATAATAAAAATATAAAAATCTCTTAAATAAAAAATTATAACAAAAAATAAAATCATAAAA 3275

Qy        610 AAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGA 669
              |||       |   |  |    || ||       ||   |      |  |||||
Db       3274 TAAATATAAATTATATCGTAACAACTAAAACCAAAACATACAAATATTAAAAAAAAACTT 3215

Qy        670 GTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATG 729
              |    ||| |  ||| |||||  ||||| | ||   |  | | | || | |||
Db       3214 TAAAAAAATTAAATTAATATTTATCGTAAATTAAAATCTAACTATAAAAACATTAAAATC 3155

Qy        730 AATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAG 789
              || | |||| |  |||||  |  |   | ||| || || ||| ||    ||| ||
Db       3154 CATTTAATTACTTCTTTAAAATAAAACTCTTTAATATATAAATATAATATATATATTT 3095

Qy        790 CCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAA 836
              | |  |   |||||| | |  || ||   ||  |   || |||||
Db       3094 ATAAAATATTCTATTCATAATACCTAACTTAAACTTTTTAATAAAAA 3048


RESULT 10
US-10-221-613-378/c
; Sequence 378, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, ALexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
;  FILE REFERENCE: 5013.1004
;  CURRENT APPLICATION NUMBER: US/10/221,613
;  CURRENT FILING DATE:  2002-09-13
;  PRIOR APPLICATION NUMBER: PCT/EP01/02945
;    DE 10013847.00
;    DE 10019058.8
;    DE 10019173.8
;    DE 10032529.7
;    DE 10043826.1
;  PRIOR FILING DATE: 2001-03-15
;    2000-03-15
;    2000-04-06
;    2000-04-07
;    2000-06-30
;    2000-09-01
;  NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 378
;   LENGTH: 8676
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (2938, 2998, 3286)
US-10-221-613-378

  Query Match             6.1%;  Score 51.6;  DB 16;  Length 8676;
  Best Local Similarity   48.9%;  Pred. No. 0.57;
  Matches  138;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy        504 ATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAA 563
              ||||  |||  |||  | |  |  | |       | |  ||   ||| | || ||  |||
Db       6360 ATAAACAAAAAAATTTTCAACAACTCATTTCTATTCTACAACCCAAAAATAAATTTAAAA 6301

Qy        564 GAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTA 623
              |||| |||  |||   |  |||||| |||     |  |||||||||    | ||  |||
Db       6300 AAAAAATTATAATTCAATTTTTCCAAATTAACATTTCCTACTTCTTAATTCTCCTAATTA 6241

Qy        624 CTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAA 683
              |||   ||    | ||  |||     ||  |||||     |   |  | | |
Db       6240 TCGAACTACCTTAAATCCTAAACTAAAACATAAAAAAAAAAAAAAAAAAAATCAATAAATA 6181

Qy        684 GAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGG 743
              | | |   | |||      | |  ||||||   | ||   ||  ||  | |||
Db       6180 TACTCTAACAAATAATACCTTACTTTCAAACAACTATATCTAAAAAAACTTTATTTAATA 6121

Qy        744 TTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTA 785
              ||  |||   | |  | |  |||||  |   ||||  |||
Db       6120 TTCCATATACATATCCTAAAAAAAAAAATAATACATATTTTA 6079


RESULT 11
US-10-425-115-647/c
; Sequence 647, Application US/10425115
```

```
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 647
;   LENGTH: 802
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(802)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_100591C.1
US-10-425-115-647

  Query Match             5.9%;  Score 50.6;  DB 18;  Length 802;
  Best Local Similarity   44.9%;  Pred. No. 0.35;
  Matches  227;  Conservative    0;  Mismatches  277;  Indels    2;  Gaps    1;

Qy        265 AGGAAGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGG 324
              || |||||    |  |    ||| |   |||        | | | || || | |
Db        745 AGAAAGAANAAAAAGAAAANAATAATTTAAAAATTAATATAAAATAAAAAANAAAAAAAA 686

Qy        325 AAGTAATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACG 384
              ||   |   | | | |    || ||  | |  |      | |    ||| ||      ||
Db        685 AAAATAATATAAAAATTAAAAAAAAAAAAAATAAAAAAAAAAAAAATAAAATAAAAAATAAA 626

Qy        385 ACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGG 444
              | | || | |||    | |||   | |||    |  |  |    | |  |||    |
Db        625 AAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAATTAAAAATACATAATTAATAAAAAAAA 566

Qy        445 TTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGA 504
              |  | |  ||||| |||  |  | |||    |||||||  ||   |  | ||   |
Db        565 AAAATATAATAATAAAATAAGACAAATATTATAGAAAGTAATAAAAAATAAGAAAAAGAA 506

Qy        505 TAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAG 564
               ||  |||  ||||   |||| | |  |||     |  || | ||||| ||||| |  |
Db        505 AAACAAAAAAAATAGAAATAAATAAAAAAAAAAAATAAAAATATAAATAAAAGTTATGCAA 446

Qy        565 AAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTAC 624
               ||||  |  |  || ||   | |    | |        || |   |||  |    ||
Db        445 TAAAGAAATAAAATTGTATGACAAGACAAATATATTAAACAACAATAAAAAAAAGAATAT 386

Qy        625 TTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAG 684
                  ||||     || | |   ||  ||| |     | | | |    || ||
Db        385 AAAAAATTATAAACAATAATAGAATAAAATAA--AAATAATAATAATAAATAAATAAAAA 328

Qy        685 AATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGT 744
              ||  |  | || |    | || |   ||| |||  ||    ||  || |   ||| |
Db        327 AACATAAAAACAATTAAAAAATTCCAAAAAACATTAAAAGTAACAAAAAAAGTTAATATA 268

Qy        745 TTAATAATAGTGTATTTTATTTAAAA 770
              ||||||  |  | |   | ||||||
Db        267 TTAATAGAAAGGAAAAAAAGTTAAAA 242


RESULT 12
US-10-311-455-1473/c
; Sequence 1473, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1473
;   LENGTH: 9157
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1473

  Query Match             5.9%;  Score 49.8;  DB 15;  Length 9157;
  Best Local Similarity   45.7%;  Pred. No. 1.4;
  Matches  174;  Conservative    0;  Mismatches  207;  Indels    0;  Gaps    0;

Qy        456 TTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTA 515
              || | |||    | |||  || |  | || | |  |  |  | ||| | || |    ||
Db       6328 TTTATAAATTAAATTTCAAACGTACATTATTTACATAAAAAACTACTAACAAATCGTATA 6269

Qy        516 ATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCA 575
              ||     | |  ||  | | |       || | ||||| || |   ||| | || |||  |
Db       6268 ATTCCGTTTATATAAAATATTCAAAATAACAAAAATATAAATACAAAAAATAAATTAATA 6209

Qy        576 TTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACA 635
              || | |    | ||  |||    |  ||  || ||||  |     ||| ||||||  | |
Db       6208 ATTACCAAAAACTACAAATTAAAAATAACAACTAAAAAAAATAAAATACGTTTTTAAAAA 6149

Qy        636 AACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATAC 695
              |    | |    ||      |    ||    |      ||  ||  || |   |  |  |
Db       6148 TATAAATAATTTTAAAATTTATTCTAATAATAACTACACAACTCTCAATATCTTAAAATC 6089

Qy        696 TAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGT 755
              |    |  |   ||  | | | | |      |   | |||||   |  |   ||   |
Db       6088 CATTAAATTATATACCTAAATAAATAAATTCTATATATTAATTATATCTCAATAAAACTT 6029

Qy        756 GTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCACAACTC 815
              | |||| || ||||| |    ||||   | ||      |  | || ||    | | |
Db       6028 TTTTTTTTTTAAAAAAAAAAAAATACTAAAACATATTTAAAATATTCTTTCAAAAACTTA 5969

Qy        816 ACAACAAAGTACAAAGAAAAA 836
              | | ||  | |||  |||||
Db       5968 AACAAAACTAAAAAATAAAAA 5948


RESULT 13
US-10-311-455-1701/c
; Sequence 1701, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
```

```
; SEQ ID NO 1701
;   LENGTH: 17721
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1701

  Query Match             5.8%;  Score 49.6;  DB 15;  Length 17721;
  Best Local Similarity   47.2%;  Pred. No. 2.1;
  Matches  151;  Conservative    0;  Mismatches  169;  Indels    0;  Gaps    0;

Qy        510 AAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAG 569
              ||| |||| |  ||||   |||    |  ||  | || ||||| ||       |   |||
Db      14890 AAAATAATTAAAATAACAAATTTTTTTACTTCAATAAAAAATAAAACACACTATATAAAA 14831

Qy        570 TTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTT 629
                ||     |  ||||| | |  ||  |    | ||  ||| ||     | | |||
Db      14830 AATTAATAAAACAAATCCAAATAAAATCCAATTTATACTATAAAATAAAATTAAATTTAA 14771

Qy        630 TACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGT 689
              || || ||    |    ||| | |   || |      | | | |||  | || | |
Db      14770 TATACTAAAAAATATTTTAAATTTAATATAATATTAATCAATAAATTCCAAAAAAATTTA 14711

Qy        690 TTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAAT 749
              |||||  |      |||| | ||    |    | |  ||      |  || || ||   |
Db      14710 TTATAAAATTCACTAAATTATAATATTACTTTTTATCATTTTCCTATTTTTTTTTTATTT 14651

Qy        750 AATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCA 809
               ||  | | ||  | || ||| | |    |   |  ||||| ||| |  ||  ||
Db      14650 TATTTTATTTTCTCTTAAAACAAAAATCTCGCTATACCACCCAAACTAAAATACAATAAC 14591

Qy        810 CAACTCACAACAAAGTACAA 829
                | || | ||  | |||||
Db      14590 ATAATCTCGACTCACTACAA 14571


RESULT 14
US-10-425-115-160787/c
; Sequence 160787, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 160787
;   LENGTH: 578
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(578)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_78218C.1
US-10-425-115-160787

  Query Match             5.8%;  Score 49.4;  DB 18;  Length 578;
  Best Local Similarity   53.2%;  Pred. No. 0.56;
  Matches  124;  Conservative    0;  Mismatches  108;  Indels    1;  Gaps    1;

Qy        480 AAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTT 539
              ||  || |||   |  || ||   ||||  ||| | || |  ||||   | ||||  |
Db        359 AATAAAAGAAAAAGTATATAAAATATAAAAAAAATTATGAAAATAAAAAAATAAAAAGAA 300

Qy        540 TGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTC 599
                                                                        ||||||||| | | |   ||| |||||| ||||||  | | |  |||  | | | |
Db        299 ATTAAAAGAAA-AAATTTAAAAAAAAAAAGTAAGCATTAACAAATGGCAATATTTATATA 241

Qy        600 ACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCA 659
              |  ||| |  ||||     ||  ||   ||| |||  || |||| |   |||    ||  |
Db        240 ATAACTACAAAAAAAACATTGATAAAGTTTAACAGGAATGGGTAACANAAAAAAAAAAAA 181

Qy        660 AAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAA 712
              ||||     |   |       | || ||      | |  |||  | |||  | ||
Db        180 AAAAAAACAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 128


RESULT 15
US-10-311-455-2148/c
; Sequence 2148, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2148
;   LENGTH: 113515
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2148

  Query Match             5.8%;  Score 49.2;  DB 15;  Length 113515;
  Best Local Similarity   46.7%;  Pred. No. 5.6;
  Matches  156;  Conservative    0;  Mismatches  178;  Indels    0;  Gaps    0;

Qy        499 AACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTT 558
              ||||   |  | || |||||   |||    |  |     ||   | |  | ||  ||
Db      13895 AACTACCATATTAAATAATATACATATAAAACAATTCCATTCTTACATTATATTAAACAC 13836

Qy        559 GTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACT 618
                  |    | ||  ||   |||||  |    || ||| |       || ||  || |
Db      13835 CACTAATCTAAATAATATAAACTACTCCTTTTTTCTTTTTTCTACATTTTTTTAATTCAA 13776

Qy        619 TCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCT 678
                ||   |    |  ||      ||  |||    ||  | |||  |  | ||| |
Db      13775 CAAAACAACTAATAAATAAATAAATAAAAAAATAATAAAAATAAACTTAAAATTAATCAA 13716

Qy        679 TATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATT 738
              | || | |   |  |||||||  | |||  |||||   | |||  | ||  |      |
Db      13715 AACAATATTCAATTAACTAAATTATAAAATACAAAACAACAAAAAAAAAAAAACATCTTTA 13656

Qy        739 ATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAA 798
              |||    |   |||   | ||||| | | |||||  ||  || | |  | ||  || | ||
Db      13655 ATTAAATCCTTAAATATATATTTCAATCAAAAAAATAACAAAAAAAAACAAACACAAAAA 13596

Qy        799 GTTATTCATCACAACTCACAACAAAGTACAAAGA 832
                ||    | |||   | ||  | |  | ||| |
Db      13595 ACTAACATTTACATTACCCACAATATAAAAAAAA 13562


Search completed: December  7, 2004, 06:11:21
Job time : 509.824 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:35:24 ; Search time 3211.92 Seconds
(without alignments)
9654.718 Million cell updates/sec

Title: US-09-806-197-2
Perfect score: 851
Sequence: 1 ctgaaaaggaagttcatcga.....aaaaagttgcaattctgtca 851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	83.4	9.8	1101	9	CNS0039G	AL063921 Drosophil
c 2	65.6	7.7	938	9	CNS006TJ	AL065906 Drosophil
3	64	7.5	987	9	CNS014PQ	AL104456 Drosophil
4	63.8	7.5	1001	9	CNS0155H	AL105023 Drosophil
5	63.6	7.5	1092	9	CNS020K7	AL175696 Tetraodon
6	63.4	7.5	1101	9	CNS00KB5	AL077453 Drosophil
c 7	62.8	7.4	1201	9	CNS0167M	AL106396 Drosophil
8	62.6	7.4	1101	9	CNS00KAE	AL077628 Drosophil
c 9	61.8	7.3	1380	9	AG321083	AG321083 Mus muscu
c 10	61.2	7.2	1101	9	CNS016LI	AL106896 Drosophil
c 11	60.8	7.1	1101	9	CNS0039L	AL063926 Drosophil
12	60.8	7.1	1101	9	CNS0182P	AL108811 Drosophil
c 13	60.6	7.1	1101	9	CNS000B8	AL063632 Drosophil
c 14	60.2	7.1	1101	9	CNS0039R	AL063932 Drosophil
15	60.2	7.1	1101	9	CNS00EVL	AL069706 Drosophil
16	60	7.1	1146	9	CNS021G2	AL176843 Tetraodon
17	59.8	7.0	1101	9	CNS0021J	AL061936 Drosophil
18	59.2	7.0	903	9	CNS00I1C	AL074609 Drosophil
c 19	59.2	7.0	1036	9	CNS00599	AL057797 Drosophil
20	59	6.9	1101	9	CNS00EO7	AL069440 Drosophil
21	58.8	6.9	1101	9	CNS000D1	AL065414 Drosophil
c 22	58.2	6.8	1101	9	CNS001FB	AL060732 Drosophil
23	58	6.8	1200	9	CNS016CO	AL106578 Drosophil
24	57.8	6.8	928	9	CNS00DKY	AL071865 Drosophil
c 25	57.8	6.8	1101	9	CNS00EVL	AL069706 Drosophil
26	57.8	6.8	1101	9	CNS0100X	AL098379 Drosophil
27	57.6	6.8	1048	5	BX343665	BX343665 BX343665
28	57.4	6.7	603	9	CNS011I5	AL100295 Drosophil
c 29	57.4	6.7	1442	8	CC303074	CC303074 CH261-173
c 30	57.2	6.7	879	9	CNS01JRG	AL147405 Anopheles
31	57	6.7	973	9	CNS071LE	AL425064 clone BA0
c 32	57	6.7	1352	9	AG382165	AG382165 Mus muscu
33	56.8	6.7	746	8	BH921242	BH921242 odg89g08.
34	56.8	6.7	1608	9	CL118721	CL118721 ISB1-72J8
35	56.6	6.7	514	6	CA856511	CA856511 PfESToac0
c 36	56.6	6.7	928	9	CNS00DKY	AL071865 Drosophil
c 37	56.4	6.6	1135	9	AG278055	AG278055 Mus muscu
c 38	56.4	6.6	1201	9	CNS015X9	AL106023 Drosophil
c 39	56.4	6.6	1749	9	AG310735	AG310735 Mus muscu
40	56.2	6.6	1067	6	CD384866	CD384866 AGENCOURT
c 41	56.2	6.6	1101	9	CNS017KX	AL108171 Drosophil
42	56	6.6	984	9	CL113045	CL113045 ISB1-57P1
c 43	56	6.6	1805	9	CL080711	CL080711 CH216-159
c 44	55.8	6.6	974	9	CNS00ITT	AL075432 Drosophil
45	55.6	6.5	1493	9	CL078589	CL078589 CH216-151

ALIGNMENTS

RESULT 1
CNS0039G/c
LOCUS CNS0039G 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063921
VERSION AL063921.1 GI:4941778
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_lib="RPCI-98"
/note="end : TET3"
ORIGIN

Query Match 9.8%; Score 83.4; DB 9; Length 1101;
Best Local Similarity 17.2%; Pred. No. 7e-08;

Matches 111; Conservative 291; Mismatches 243; Indels 2; Gaps 1;

```
Qy        141 GATGACTTGTTGATATTTATTGAGAATTACAACTTTGCAGAAAAGCCAACTCTGCGGTTT 200
Db       1098 RRWGDDTWDRDTRKDDWDWTKWWTWWKDRADDRRWAGDADRWADDGAGTWWTATWWWWW 1039

Qy        201 CAGGTTCTGGAATTAATTAAGCTTTTACATCACTATGGAAGTGATACTATTCGAAGCGGA 260
Db       1038 WWWATWDTWWDKWWWWATAAKTDTAWTWWRTAWRADWAGRDRGAGKRDRDAATDADGAGR 979

Qy        261 GTGGAGGAAGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCT 320
Db        978 RDGGRKRKDKKDRKDGDDDKKGGKKKKAAKAAKWATKWWDDWDWDKDWKWDGAKDRKADD 919

Qy        321 CTGGAAGTAATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTT 380
Db        918 DDGAGDKDDDGKGKDADDDTDGTKDDDDKDKWDDWDKAKGTWGDATWAWAATDWWWWGWA 859

Qy        381 TACGACACAATAGAAAAGGCAGAAAGGGATAA--AGTGCTCTATGAATGGTGAGGAATTG 438
Db        858 DADWWTWDAAADDWWADDRWDAWAWKWDDAWAWGARTADRRDWGDRAGKRGGARKRRDRK 799

Qy        439 GGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAG 498
Db        798 RADDKRDAADDRDDAATWTTWTTTTTRDTDDWKWKTDTWTRWAADRTWDRDDDDDRDRAG 739

Qy        499 AACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTT 558
Db        738 TAGRKWRRTWKRRWKRRDTRWDDADADDTARDDRRRRGDDGADAGKGKKTGRKRRRRDRA 679

Qy        559 GTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACT 618
Db        678 TWDRTDAWWADAAWWTTTDTDTDDWDKRDRRRKGARRRRRTTARAAWDWWTWKAWDWAKW 619

Qy        619 TCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCT 678
Db        618 DWKTRADRWDRWAADTWTDARKADRDWAKARAWRARRDRARAARADRRWTTKGKTTTATW 559

Qy        679 TATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATT 738
Db        558 TTWAARAAWWAWWAWATTTATWTTTWTTTWTTTTTTTTTTTTWTTTAAWWWAAWWTATWAA 499

Qy        739 ATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTA 785
Db        498 WTAAAWAAAAAAAWAATTTTTTTTTTTTWAWTAAAWWTAWTWTTTTW 452
```

RESULT 2
CNS006TJ/c
LOCUS CNS006TJ 938 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC # BACR14J11 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL065906
VERSION AL065906.1 GI:4944874
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 938)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1. .938
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14J11"
/clone_lib="RPCI-98"
/note="end : T7"
ORIGIN

Query Match 7.7%; Score 65.6; DB 9; Length 938;
Best Local Similarity 22.2%; Pred. No. 0.00057;
Matches 101; Conservative 157; Mismatches 197; Indels 0; Gaps 0;

```
Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
Db        932 ADTAKAADAKAADKADAKDAGKAAADAWAARDADKKDKADKDADKAAKADDDKDARKAKA 873

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
Db        872 KKKKKKAKKKKKADKDKKAKKKAKWKAKDKWRAKKADAADAAKAKAKAAAAAWADKKAWK 813

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
Db        812 ADADKDAAAKDDKADKDADKKAKAKDDKDAKDAADADDDKAKKKDADAAKADDAADAAA 753

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTT 628
Db        752 AKAADDRADAKAGDAKAKAAAAKDKKAAAKAKDKKDDKKAAMAWADAADRDAKDCAKDKW 693

Qy        629 TTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATG 688
Db        692 WAADDKDKAAKRAAAAAAWAWAAADADDAKARTAKDAKAKAAKDKKKDAAKAAKDKKAAAR 633

Qy        689 TTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAA 748
Db        632 AGAKDKDRKDAGTAATKAAAAAAKAKKKAKDDTTDTKTWDATATDKKTDGWTKKDKATAT 573

Qy        749 TAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATC 808
Db        572 ATRGAKTGKATGKTWWTTATATAWATAAAATTAATWATATTTTTAANNAANNAAANAAAA 513

Qy        809 ACAACTCACAACAAAGTACAAAGAAAAAGTTGCAA 843
Db        512 ANANNANNANNNNAANANNNANNNANNNNNANAAA 478
```

RESULT 3
CNS014PQ
LOCUS CNS014PQ 987 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL104456
VERSION AL104456.1 GI:5616067
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 987)
AUTHORS Genoscope.
TITLE Direct Submission

```
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .987
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN12P22"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

  Query Match             7.5%;  Score 64;  DB 9;  Length 987;
  Best Local Similarity   28.7%;  Pred. No. 0.0013;
  Matches  133;  Conservative  109;  Mismatches  222;  Indels    0;  Gaps    0;

Qy        308 GTACAAGAAAGCTCTGGAAGTAATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACG 367
              | | |: |||     :||:: |::   || | |   | ||::| |  |: : |     |
Db        490 GGAWAWAAAAADRDRRGAWDAAWKAAAGAAAATTCACTYTAWKGCGTGTRTKTGTGMTAG 549

Qy        368 TCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATG 427
              : :| : :|   :  : ::    |  | |:||:: |::::::::    :: : : :   |
Db        550 WTDATRGWTGRKGGRDKDWTTAGAAAAWAADRAARDRRRRRRBKKBNKKKKKKKKTKKTT 609

Qy        428 GTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATG 487
              |   :|  :  : |           : :  :|:||:| : | : |: |  | |    |
Db        610 TTTTTKAYMWCTBKCCCCCCCCCCCCMMMAMMAMAAMAAWMAWWACWAAAATTAGAATAAA 669

Qy        488 AAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAG 547
              || :  |  | ||   |||| || :  : | :|: |:: :|:||    | |||  :|||
Db        670 AAAWAWAAAAAAAAAAATAAWTATWWAWNTNWTWTTWWWWTWTTTTTTTTTTTAAWAAAA 729

Qy        548 AAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTC 607
              ::| | || :| | |: | :: |||      : || |        |||| | :  : ||
Db        730 WWAAAWAAAWTWTTAWWATWWTTTAAAGAAAWTTATTAAATTWAAATTTTTWWTTWATTA 789

Qy        608 TTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTG 667
              | ||:  | |: ||: ||: |:: |   |     | :   | : |                |
Db        790 ATTAAWTAATTWATTWATTWATWWWAATTAATAWTWWTANAKWTTWWTTTTTTTTTWWTA 849

Qy        668 GAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAA 727
              |  |  || ::|:||  |::  : :: :|| : | ::|| : ||| | : :::  | ::
Db        850 TAAATTWTTWWWAWAAATAWRGRWTWWAWAATWAAAWWATWWAAAAWAWWWWDRAAAWDW 909

Qy        728 TGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAA 771
              : :: :| | ::||     || ||| |     |  ::::::::::
Db        910 WWWWAWTAAAWWTTAAAAAAAAAAAAWAAAAAAWWWWWWWWWW 953


RESULT 4
CNS0155H
LOCUS       CNS0155H                1001 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN13C23 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL105023
VERSION     AL105023.1  GI:5617037
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1001)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1001
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN13C23"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

  Query Match             7.5%;  Score 63.8;  DB 9;  Length 1001;
  Best Local Similarity   32.5%;  Pred. No. 0.0014;
  Matches  147;  Conservative   87;  Mismatches  218;  Indels    0;  Gaps    0;

Qy        385 ACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGG 444
              | |||| ||:::: : :|:|   || :|         |  || :| |:  ||:: |:|:
Db        547 AGACAAAAGRRVRARWWGDAGAAGAADAGAATANAGAAAAAAKGRGMAAAABDAGRASAA 606

Qy        445 TTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGA 504
              : :   :|  ||||| :: ::  |          |  :|:: ||    |   ::    |
Db        607 ARKTKWGARGGAAAAAAAGWDADDAATWMRKAARAAKADADWAAAAACAAAANVHNAAWA 666

Qy        505 TAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAG 564
              |:|:|||: :| ||  | :||: |  |||  :   ::|:|| |||||  |:     |::
Db        667 AARGKAAARRRAAAAAAAAARAGDAARAAAAWGRAAARDARAAAAAATAAAWWAATAARWA 726

Qy        565 AAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTAC 624
              |:|::  |  |    |  : |   :: : |:::   |  |    ||: |   | |    |
Db        727 ARARDARAAAAAAATTAWWTAAWWWAWAAWWWWAAATAAAAADTTWTATWAAATAAAAAA 786

Qy        625 TTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAG 684
              ::|  | | | |     : :|  |||  |  ||    ::|        ::|       || :
Db        787 AWWTAAAAAAATATTTAWAWAWAAAAATWAAAAATTWWAAAAWWTAWWAAAAAAAATTWA 846

Qy        685 AATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGT 744
              | | |||| |  :||: |  :|:    || | | |||  | |:  :|||  |||:    :
Db        847 ATTTTTTAAAWAWAAWWAATWAWATATAAAADAAAAAAAAAAWAAWATAAAATTWAAAAW 906

Qy        745 TTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATT 804
              ||::: |:|  :: :|||:||: ::|| |    |:    :   |  | :  |  :||
Db        907 TTWWWTAWAWADWTWTTTWTTWTWWAAWAAATAAWTWTAWTTAATWTAAWAATAAAWATA 966

Qy        805 CATCACAACTCACAACAAAGTACAAAGAAAAA 836
              |  | :|   : |:  || :  ||| ||||:
Db        967 AAAAATWAWAAWWAWATAAWWTAAAAAAAAAW 998


RESULT 5
CNS020K7
LOCUS       CNS020K7                1092 bp    DNA     linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            222L11 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL175696
VERSION     AL175696.1  GI:7813753
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetradontoidea; Tetraodontidae; Tetraodon.

REFERENCE 1
AUTHORS Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645

REFERENCE 2
AUTHORS Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143

REFERENCE 3 (bases 1 to 1092)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.

FEATURES Location/Qualifiers
source 1. .1092
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="222L11"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG222CF06LP1~end : T7"

ORIGIN

Query Match 7.5%; Score 63.6; DB 9; Length 1092;
Best Local Similarity 38.8%; Pred. No. 0.0016;
Matches 156; Conservative 54; Mismatches 190; Indels 2; Gaps 1;

```
Qy        445 TTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGA 504
              ||||  :::    |||||||   : || ||   : ||  ||: |: :  : :  ||   |
Db        622 TTTAAWWWWAAAAAAAAAAAAAWWTTTTTTTTAWAAAAAAAWAAWWWWWWWTTAAAAAA 681

Qy        505 TAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAG 564
               || |    |  |  || |   |||  |||        |||: :|||: || :   |||
Db        682 AAAATTTTTTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAWTWAATWAAAWWAAAAAAW 741

Qy        565 AAAAGTT--AGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTT 622
              ||:| ::  :  :|||| |: ::  ::   | :    :  : |  ||  |: :| :  :|
Db        742 AAWAAWWWAWTTWTTTTTTWAWWWWWWWAAAAWAAAWWWTWTTAWTTTTAWWWATWAWWT 801

Qy        623 ACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATA 682
               | ||||| | |||    |  |   |:    ||  :|||   |    | |  : :| ||
Db        802 TWTATTTTAAAWAAAANAWTTTAWTTAWTAAAAAAWWAAATAATTWTTTAAAAWAWTTTA 861

Qy        683 AGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTG 742
              : ||| :: ||:  |||  | |||| | ||| : |  ::|     | : :|:  :||
Db        862 WAAATTWWAATWTAAAAAAAAAAATTAAAAAAWAAWTWWTTTTTTTTWTWAWWTWTAAAT 921

Qy        743 GTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTA 802
               |||  |  |  | ::|||| :| |:||| |    | |   ||  :   | ::  | ||
Db        922 WTTTTTTTTTTTTTWWTTTTTWTAAWAAAAAAAAAAAAATATAAAWNAAAAWWTTAATTT 981

Qy        803 TTCATCACAACTCACAACAAAGTACAAAGAAAAAGTTGCAAT 844
               | | | ||   | || |||  | | | |||:|  |  |||
Db        982 ATTTTTAAAAAAAAAAAAAAATAAAATATAAAWAAATAAAAT 1023
```

RESULT 6
CNS00KB5
LOCUS CNS00KB5 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC: BACR39P05 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL077453
VERSION AL077453.1 GI:4956930
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR39P05"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

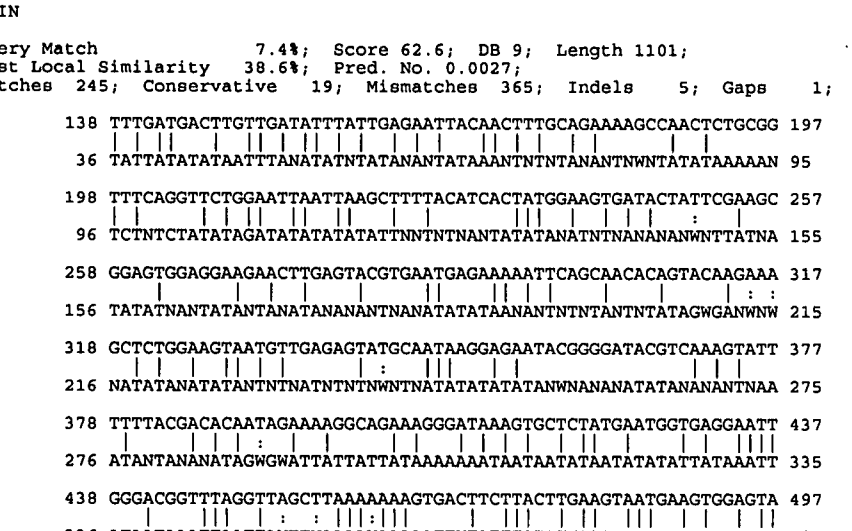
Query Match 7.5%; Score 63.4; DB 9; Length 1101;
Best Local Similarity 42.4%; Pred. No. 0.0018;
Matches 236; Conservative 11; Mismatches 306; Indels 3; Gaps 1;

```
Qy        277 AGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGTAATGTTGA 336
              | |    ||  | |||||    |  || | |  | || |||  |:   ||  ||     |
Db         31 ANAANNANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 90

Qy        337 GAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAA 396
              |  |   || ||  | || |      | |    ||| ||      |  | | || | |||
Db         91 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 150

Qy        397 AGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCT 456
              |   | |||       ||       ||| ||   |    |: |   |    | |  ::|
Db        151 AATAATAAATAATATAAAAAAATTATAAAATATTTAAAWATWATAAAAATAAATWWATTA 210

Qy        457 TAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAA 516
              ||  |::  |    : | |  ||  ||:  |    | :| |: | : ||  |||  ||
Db        211 AAATTAWWAAAAWAAAWAATTAAAAAAAAWATAAAAAAAWAAAWATAWAAAAAAAATAAA 270

Qy        517 TAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCAT 576
              ||  | ||   | ||||     |  |||| ||    ||     ||| |||| |
Db        271 AAAAAAAAAAAAAATAAAAAAAATAAAAAAAAAATATAAAAAAAAAAAAAAAAAAAANANN 330

Qy        577 TTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAA 636
                | |   ||   | || |    | ||  |||   || || ||  ||||||
```

```
Db        331 NNNNAAATAAAAAAAANAATTATANANAATTTATAA---TATTTTTTTATTTTTTTATTTT 387

Qy        637 ACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACT 696

Db        388 ATATTTTAAAAAAATAAAAATAAAAAAAAAAAAAATAAAAAAAAATAAAAAAAAAAAWATA 447

Qy        697 AAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGTG 756

Db        448 AAAAAAAAAAAAAAAATNAAANAAANAANNNNAAAAAANAAATATAATTTAAAAAAAATT 507

Qy        757 TATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCACAACTCA 816

Db        508 TTTTTTNTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAATAANAAAAAAN 567

Qy        817 CAACAAAGTACAAAGA 832

Db        568 ANAAANAAAAAAAANA 583


RESULT 7
CNS0167M/c
LOCUS       CNS0167M                1201 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN15M24 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106396
VERSION     AL106396.1  GI:5621701
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN15M24"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
ORIGIN

  Query Match             7.4%;  Score 62.8;  DB 9;  Length 1201;
  Best Local Similarity   35.7%;  Pred. No. 0.0024;
  Matches  200;  Conservative   63;  Mismatches  297;  Indels    0;  Gaps    0;

Qy        285 AATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGTAATGTTGAGAGTATGC 344

Db       1201 AWAWATWAWATAWAWATATATAAATATAAAAATAAATAWAWAATATATAWNAAAAANTAT 1142

Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404

Db       1141 AAAAAAWAAAATAWWATATAAAWAAWAAAWWWWATATTAAATAWAATATAANWAAAAAAA 1082

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAA 464

Db       1081 AAAAAAAAAWWWTTTHTANAATATTTWTTTWNTATAWAATWTTTTTTTTTTTTTTTATA 1022

Qy        465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524

Db       1021 WAWAAAWAAAAAAAAATTTTAAAAATAAAWTAATTATWAAAATTTTTAAAAATTTTTWAT 962

Qy        525 AGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTCTACT 584

Db        961 WWTTTTTWAAAAAAAAAAWAATATWAAAWTTTTTTTTATWTATAAAWAWAWTTTTTTYWT 902

Qy        585 TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCA 644

Db        901 TAAAAAAAAAAWTTAAAAATWTAAAAATTATAAAATAAAAWAAAAAAAAAAAAAAAAWW 842

Qy        645 AGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGA 704

Db        841 AAAAAWTTTWATTATAAWATTTTTWAAAAAAAAAAAATTTAATTYTTTTTNAAAATAAAA 782

Qy        705 AATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATTTTAT 764

Db        781 AAWAAAWAWWWAAAAAATWWTAAATATAAWTTAAATNCATAAAACAAAAAAWAATTWHAT 722

Qy        765 TTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAG 824

Db        721 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATTTTTTTHACAAMMAMMAMMYMMMMCAAAA 662

Qy        825 TACAAAGAAAAAGTTGCAAT 844

Db        661 MAMMAAVMAAAMACMAMMMT 642


RESULT 8
CNS00KAE
LOCUS       CNS00KAE                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR39P05 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL077628
VERSION     AL077628.1  GI:4956903
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR39P05"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
```

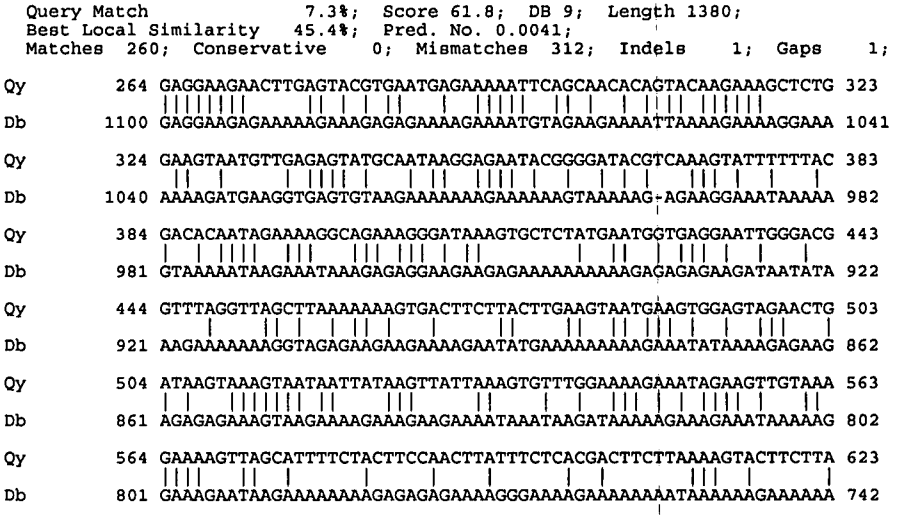
ORIGIN

Query Match 7.4%; Score 62.6; DB 9; Length 1101;
Best Local Similarity 38.6%; Pred. No. 0.0027;
Matches 245; Conservative 19; Mismatches 365; Indels 5; Gaps 1;



RESULT 9
AG321083/c
LOCUS AG321083 1380 bp DNA linear GSS 02-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-107D08.TJ, genomic survey sequence.
ACCESSION AG321083
VERSION AG321083.1 GI:47894040
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1380)
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institude, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES Location/Qualifiers
source 1. .1380
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-107D08.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

Query Match 7.3%; Score 61.8; DB 9; Length 1380;
Best Local Similarity 45.4%; Pred. No. 0.0041;
Matches 260; Conservative 0; Mismatches 312; Indels 1; Gaps 1;



RESULT 10

```
CNS016LI/c
LOCUS       CNS016LI                1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106896
VERSION     AL106896.1  GI:5624374
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN16D22"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
ORIGIN

  Query Match             7.2%;  Score 61.2;  DB 9;  Length 1101;
  Best Local Similarity   32.6%;  Pred. No. 0.0054;
  Matches  107;  Conservative   76;  Mismatches  145;  Indels    0;  Gaps    0;

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              |:|  :::|  |   |:   |  ::|   :    :||: :| |   |  || : || : |:
Db       1077 TMAYWWWWTYCTMTMAMTAMAAAWMAAAMWMAAWMAAWAWATTTAYATTTHWTAATHTAW 1018

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTT 628
               :::  || ::  : ||   | : ||:| | :  |::| |:: :  :: |:::|: ::|
Db       1017 TWYWCAATAHWWAWWTTTATATWAATHTATTWATAHWTATYWTWTWHWATWYWTMTYWTA 958

Qy        629 TTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATG 688
              |:| |:|||      ||  ::|    :|  :| ||  :  : :|| :| |:|: |  ||
Db        957 TWAWAHAAAWATASTAAAYWHATAATWAWAWATAATTWTWWTWTAWWTWTWAWTAATATW 898

Qy        689 TTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAA 748
               |:|  :  | :: : ::|: : | | : : |:   :  || :  || ::  |:  |||
Db        897 ATWAAMTAATWWWWAWHAWWWTAYAWWWWWAWTATWWTATAWTAATAWWATTWTAATTAT 838

Qy        749 TAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCATC 808
               : :|   ||| | :|: ||||| |:  :| : |: |  |   | |  ||  ||   :
Db        837 AWTWAAAATATATAWTWAAAAAAAAWATHAAWAAWAAATATWAAAAAAAAHMTAAAAWAT 778

Qy        809 ACAACTCACAACAAAGTACAAAGAAAAA 836
              | ||:  | || |||  | ||| |||:|
Db        777 AAAAVAAAAAAWAAAWAAAAAAAAAAHA 750


RESULT 11
CNS0039L/c
LOCUS       CNS0039L                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR08I10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063926
VERSION     AL063926.1  GI:4941783
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08I10"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

  Query Match             7.1%;  Score 60.8;  DB 9;  Length 1101;
  Best Local Similarity   30.4%;  Pred. No. 0.0066;
  Matches  127;  Conservative  109;  Mismatches  177;  Indels    5;  Gaps    1;

Qy        151 TGATATTTATTGAGAATTACAACTTTGCAGAAAAGCCAACTCTGCGGTTTCAGGTTCTGG 210
              |: | :|:: :|||::  : :   :|  |||:||:  || : |      :      : : :
Db       1099 TDKTGKTKWRWGAGDRRADRWGTRKTWRAGARAARRAAAAWRTAWAWAWAATAAAKWKWR 1040

Qy        211 AATTAATTAAGCTTTTACATCACTATGGAAGTGATACTATTCGAAGCGGAGTGGAGGAAG 270
              :|: ||::::| | :|: |:    ::::||: :|::  | :  |:: :: |: ||
Db       1039 WAWRAAWWDWGATGDTWTADWGDADDRRAADRRADRRRAADATARKRKKTGWTGA----- 985

Qy        271 AACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGTAA 330
              :|  |: :|:  | :|: || :|  :   : :       ||  |     |   ::|  |:
Db        984 RATRTKTRTRWTTWRAWAAGTRATKWGDTDARTWTGTTTARTATGKKGTATWKGTTGKTT 925

Qy        331 TGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAA 390
              ::||::|  :::   |:|:  | ||| |      |:|  | | |  ||:|:: |    | |
Db        924 KKTTKRGGTKRKTKKAKAWAAAAAAAAAAAAAAAWAWWTTATATAATDTKWRATAGWAWAT 865

Qy        391 TAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTAGG 450
              |  :| |:  : :|  ::  || :    : :|  ||     :| |||  : :|      |:
Db        864 TKADATADWAWADATKRTKTATRAATTDKDAAAAAAAAAARATGAAAARARAAAAAARAKW 805

Qy        451 TTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTA 510
              |:::   |:::|||:: : |  | | :    ::|| :: |   |  |:|:   | || |:
Db        804 TDDRWAAARWDAAARDARGTGGTAARDGWGGKKAAAKWTGRTTAAAADAWAAAAAAAATW 745

Qy        511 AAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              |: :|||||:|| ||  || : :|   |:: :|||| |:|:| |:  ::|||| ||||
Db        744 AWWWAATAAKTAAAAAATARWGWAAAATKWTKAAAAAADAWAAARAGDKTAAAAAAAA 687


RESULT 12
```

```
CNS0182P
LOCUS       CNS0182P                1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL108811
VERSION     AL108811.1  GI:5629115
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN37D10"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

  Query Match             7.1%;  Score 60.8;  DB 9;  Length 1101;
  Best Local Similarity   20.6%;  Pred. No. 0.0066;
  Matches  102;  Conservative  195;  Mismatches  193;  Indels    4;  Gaps    1;

Qy        285 AATGAGAAAAATTCAGCAACACAGTACAAGAAAGCTCTGGAAGTAATGTTGAGAGTATGC 344
              |:: |:: :||   :  || | |  | || ||   : |  ||  |  |: : | |:| :
Db        604 ARWAAKWTWAAAAARAAAAAAAAAAAAAAAAAAAGTTKTTAAAAAAAGGGWGKTGKGKAADG 663

Qy        345 AATAAGGAGAATACGGGGATACGTCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAA 404
              :  ||  : ::   |:|:: :    ||| :|:     |  | | || | ||:||| :::
Db        664 DGGAAAADAWRRGRGRGRRARAARAAAAADAKRAAAAAAAAAAAAAAAAAAAADAGGKRKWK 723

Qy        405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGT----TTAGGTTAGCTTAAA 460
              |:| :    |:   :|:|: : ||    : ::   |   :    |:|::  ||  :::|
Db        724 KGDGKDGGKATKTAAWAKGRKDWGTATAWWTWDATWADTWKAATDTDAKRAAAGRRKRDA 783

Qy        461 AAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAAT 520
              :  |::|     | :  : | |::|    |:  :|  : |:  :::|:  |||  || ||
Db        784 RKTARDGGRRARTRRRAWAAGGKRARAGARRARAARRAADDRDWDAWAAAAAAAAAAAAA 843

Qy        521 TATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAAGTTAGCATTTTC 580
              : |::: ::::::::|:::::  ||: ::| |:|::  : ::::::|::  :: |::|:
Db        844 WTTWRDRWWDWDDWDTRWDDDTTAAWWDDARARARRRRRRRRRRRRRARRAARRAADDTDT 903

Qy        581 TACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACACAAACGG 640
              :: :      :: |:| | : :: :: : ::::    :  ::: :   ::: | | : :|
Db        904 KDRWADATTDKDTTKTWTTDDDDWDKAKRDRWWAAKADGAWKWRDRARDDWAATAKDDDG 963

Qy        641 GTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATACTAAAT 700
              :  :||:  : |:: : ::      :::::::: |:  ::::  :|:::::::  : :
Db        964 WKDKWGGRGKRGGKDKKRWDKGKTGKKDDDDWDKTWTRDWWWWWTRTKWDWWWDDGRGRG 1023

Qy        701 GAGAAATGACAAACACAGAAATGAGAATGAATATGATTATTGGTTTAATAATAGTGTATT 760
              :: : ::|| ::| | |::: :| : :| : :::: :| : :|:::: :: : : :: :
Db       1024 RWTRKRWGAWWRADAWARDDTDGKDTRTADKDRKRTDTTKRDGDDWRKDRKDKKRRDKKG 1083

Qy        761 TTATTTAAAAAGAT 774
              :: |:::| :::::
Db       1084 DDKTKKDATWDDDD 1097


RESULT 13
CNS000B8/c
LOCUS       CNS000B8                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR01A24 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063632
VERSION     AL063632.1  GI:4938680
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR01A24"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             7.1%;  Score 60.6;  DB 9;  Length 1101;
  Best Local Similarity   38.3%;  Pred. No. 0.0073;
  Matches  171;  Conservative   56;  Mismatches  218;  Indels    2;  Gaps    1;

Qy        389 AATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGAATTGGGACGGTTTA 448
              :|:|  |||   | ||:   || || |  | ::  ||      |  | |     |  : |:
Db       1099 WAWAATAAAAAAAWAAWAAAATTAAATATTTWWAAAAAAAAAAAAAATTATAAAWATWATW 1040

Qy        449 GGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAG 508
                : :   |||||||| | |  | :|| |  ||  | :  | :  | || :  | ||||
Db       1039 TWWAWTAATAAAAAAAATAAAATTWTAATATAAAAATWWTAAWATAATATWTTTAATAAA 980

Qy        509 TAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGAAGTTGTAAAGAAAA 568
              :  |  :|||: |:||:   |  |||    |  |||| ||: || |  : |: | ::|
Db        979 WTTAAAWATAWATWTAWAAAAAAAAAAAAAATAAAAAAAAAAWAAATAAAWTTWTAAWWAT 920

Qy        569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTT 628
              |    |   |   :   :   | |:   |       |:: || |     ||  ||||
Db        919 WAAAWATATAATTTAAWWATWAAAAATWAAAAWTTTAAWTWWTAAATTTAATTTTTTTT 860

Qy        629 TTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATG 688
              :  | |||    | ||  ||| |  |   ||||:     :  ||  :  :|||| ||
```

```
Db        859 AWTAAAAAAAAAAATWAAAAAAAATAAATWAAAAAWWAAAWWWATAWAWAAWATAAWAAAT 800

Qy        689 TTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATATGATTAT--TGGTTT 746

Db        799 WAAAAAAHTAAATWAAAATTACMAAAATYWAATTYAAAATAAAWWAAAWTWTWAWAWAWW 740

Qy        747 AATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTCA 806

Db        739 WATAWAWAAAAAATTAWATTHCMAAATWAAWCAAAAAHAAWAMCAMAAAAAAATATTTTY 680

Qy        807 TCACAACTCACAACAAAGTACAAAGAA 833

Db        679 CCCAAWATCACAAYMACCAACACAAAA 653


RESULT 14
CNS0039R/c
LOCUS       CNS0039R                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08O10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063932
VERSION     AL063932.1  GI:4941789
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08O10"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             7.1%;  Score 60.2;  DB 9;  Length 1101;
  Best Local Similarity   31.7%;  Pred. No. 0.009;
  Matches  217;  Conservative  128;  Mismatches  333;  Indels    6;  Gaps    2;

Qy        135 TTATTTGATGACTTGTTGATATTTATTGAGAATTACAACTTTGCAGAAAAGCCAACTCTG 194

Db       1037 TDADAWGAKAATKWAGADAWAKATAKAGAKATAKAGWRTKWRADWDAATAKAKAKADRAG 978

Qy        195 CGGTTTCAGGTTCTGGAATTAATTAAGCTTTTACATCACTATGGAAGTGATACTATTCGA 254

Db        977 ADGAWRTAGTGTGWNKAGTGWADRTATAKATAKAGATARAAKGTRWGTRRTATAWRTATA 918

Qy        255 AGCGGAGTGGAGGAAGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAG 314

Db        917 TKARTA----TAAWWWADWTDGKDKWGTARWATRTAADAATATTKWTATARGGGRRAKRT 862

Qy        315 AAAGCTCTGGAAGTAATGTTGAGAGTATGCAATAAGGAGAATACGGGGATACGTCAAAGT 374

Db        861 ATAKAKATAKAGADTATAADKKAWATAKAWADAWGTAKGTATATATKWDGTAKRKGRGRT 802

Qy        375 ATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATGGTGAGGA 434

Db        801 GAGTAAADAGARAAARGKRWRDAAAAAAAKGAGAGAAATAAAAAAKAWARATATGGRKRA 742

Qy        435 ATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGA 494

Db        741 GDAAGAGRGGADGKGGGKGGTGAAAAAAAAAAAAKAAAKTDTAWAAAAATAAAAAAAAAA 682

Qy        495 GTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAGAAATAGA 554

Db        681 AAAAAAAAAAAAAAAARAATKRKAKAKARAKAAWKAKAKAWDAKRARTAAAAAAAAAAGA 622

Qy        555 AGTTGTAAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTAAAAG 614

Db        621 KTAAKWAAAAAAAADAAAAAAAAAAAAAAAAAAAAAGAAARTGWTAWGWTGAKNGGAAAAA 562

Qy        615 TACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGCTGGAGTTAC 674

Db        561 AAAAAAAAAKTGAWTAWGAAAAAAAWDAAAAAAAAAAAGKAA--AAAARAAAAAAAAAA 504

Qy        675 TTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAGAATGAATAT 734

Db        503 AAAWAWAWATAWTTTTATTTATRWTTWTAAAWTWAAAWAAAWAAAAAWAATWWWTGAAAA 444

Qy        735 GATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGA 794

Db        443 AWTWAATTAATTAWAAATAWWWWAAGRTAAAAAAAAAAAWAAAAAWWAATAWAWAAWTT 384

Qy        795 TGAAGTTATTCATCACAACTCACA 818

Db        383 AARATTGRATNCTAATHACTYAAA 360


RESULT 15
CNS00EVL
LOCUS       CNS00EVL                1101 bp    DNA     linear   GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION     AL069706.1  GI:4949849
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
```

```
           filters for hybridization from the BACPAC Resource Center can be
           found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR29B23"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

  Query Match             7.1%;  Score 60.2;  DB 9;  Length 1101;
  Best Local Similarity   34.2%;  Pred. No. 0.009;
  Matches  161;  Conservative   80;  Mismatches  228;  Indels    2;  Gaps    1;

Qy        368 TCAAAGTATTTTTTACGACACAATAGAAAAGGCAGAAAGGGATAAAGTGCTCTATGAATG 427
              |::::  ::  |:|:  | : ||:  | :|   | |||   :::|  |  : ::: :||
Db        463 TMMMMMMWWAATWTWWAAAWAAAWTTATWAATWAAAAAWAWWWWATTTTTWWWWWTWATT 522

Qy        428 GTGAGGAATTGGGACGGTTTAGGTTAGCTTAAAAAAAGTGACTTCTTACTTGAAGTAATG 487
               |    ::|:          |  :    || |||::| | : ||  :| || |:  ::|
Db        523 WTTTWAWWTWTAWTAAAAAAAAAWATAATTTAAAWWAATAWATTAAWAATTTAWAAWWTA 582

Qy        488 AAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTTTGGAAAAG 547
              | |    || || |  :|   | |  || || |||   ||: |||| | |||   ||
Db        583 TATTAATWTATAAATWTWATTAATATAAAAAAATATTTTTTWATAAAATTTTTAATAATT 642

Qy        548 AAATAGAAGTTGTAAAGAAAAGTTAGC--ATTTTCTACTTCCAACTTATTTCTCACGACT 605
              |||  :   |  ||  |: : ||:    |||   || ||  || | :      |  |
Db        643 TAATTAWTTATTAAATAAWTWATTWWWTAATTAAATAATTTWAAATAWAAAAAAAAAAAA 702

Qy        606 TCTTAAAAGTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAGCAAAAAGC 665
             :| :| :| |: | : |  :||| | :||      ||  | : :   |  | ::|
Db        703 AAAWATWAAWAATWATAWATAAWTTAAAAWAATAAAAWAAWAATWAWWATAATATWWATA 762

Qy        666 TGGAGTTACTTCTTATAAGAATGTTTATACTAAATGAGAAATGACAAACACAGAAATGAG 725
              |  | ||:  : :  ::| ::| | :|||  | |: | |:|:   ||  :   :| : |
Db        763 TATAWTTWTAWWWATWWAWWWTATAWATAWAATAWAAWAWAWATAAATAWATAWATWAAA 822

Qy        726 AATGAATATGATTATTGGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTA 785
              :|: :|||: ||:||   :: |  :|:|    ||:| || |:|| :      | | || :
Db        823 WAWAWATAWWATWATATAWWAATAWAWAAAAAATWTAATATWAATWATAAWAAAAAATAW 882

Qy        786 CCAGCCAGATGAAGTTATTCATCACAACTCACAACAAAGTACAAAGAAAAA 836
                  |  :  :: |    | |: | : || |||  | ||| || ||
Db        883 AWTTWWTWTTTTWWAWWATATAAAWAWATAWAAAWAAAAAAAAAAAAATAA 933


Search completed: December  7, 2004, 05:42:11
Job time : 3214.92 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 16:35:35 ; Search time 138.012 Seconds
(without alignments)
10279.455 Million cell updates/sec

Title: US-09-806-197-24
Perfect score: 30
Sequence: 1 aacaatgtatgtccggtgtacatctatgac 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30	100.0	2048	6	E40089	E40089 Plant promo
	2	30	100.0	2048	6	E40090	E40090 Plant promo
	3	30	100.0	2052	6	E40087	E40087 Plant promo
	4	30	100.0	2052	6	E40093	E40093 Plant promo
	5	30	100.0	2056	6	E40091	E40091 Plant promo
	6	23.6	78.7	2042	6	AR076817	AR076817 Sequence
	7	23.6	78.7	2042	6	E15125	E15125 Promoter. 7
	8	21.8	72.7	348283	2	AC125706	AC125706 Rattus no
c	9	21.2	70.7	161762	2	CR352244	CR352244 Danio rer
	10	21.2	70.7	213033	2	AC110515	AC110515 Mus muscu
c	11	21	70.0	140142	8	AC146937	AC146937 Oryza sat
	12	21	70.0	185386	10	AC025586	AC025586 Genomic s
	13	21	70.0	205668	10	AL592422	AL592422 Mouse DNA
	14	20.6	68.7	69674	9	AL136130	AL136130 Human DNA
	15	20.6	68.7	78026	2	AC023516	AC023516 Homo sapi
	16	20.6	68.7	215896	2	AC046180	AC046180 Homo sapi
c	17	20.4	68.0	118001	9	AC073081	AC073081 Homo sapi
c	18	20.4	68.0	193910	2	AC040996	AC040996 Homo sapi
c	19	20.4	68.0	194235	9	AC109822	AC109822 Homo sapi
c	20	20.4	68.0	288385	2	AC092430	AC092430 Homo sapi
	21	20.2	67.3	121101	2	AC008034	AC008034 Homo sapi
	22	20.2	67.3	155313	2	AC068315	AC068315 Homo sapi
	23	20.2	67.3	157007	2	AC066599	AC066599 Homo sapi
c	24	20.2	67.3	161903	2	AC021996	AC021996 Homo sapi
c	25	20.2	67.3	185067	9	AC022382	AC022382 Homo sapi
c	26	20.2	67.3	189430	2	AC011610	AC011610 Homo sapi
c	27	20.2	67.3	197360	2	AC018829	AC018829 Homo sapi
	28	20.2	67.3	216530	2	BX649535	BX649535 Danio rer
	29	20.2	67.3	260792	2	AC125770	AC125770 Rattus no
c	30	20.2	67.3	320398	2	AC136192	AC136192 Rattus no
	31	20	66.7	81398	2	AC023821	AC023821 Homo sapi
	32	20	66.7	103152	2	AP004091	AP004091 Oryza sat
	33	20	66.7	142010	2	AC139637	AC139637 Mus muscu
	34	20	66.7	158601	8	OSJN00059	AL606622 Oryza sat
	35	20	66.7	163328	8	AP004865	AP004865 Oryza sat
c	36	20	66.7	172879	10	AC142474	AC142474 Mus muscu
	37	20	66.7	201294	2	AC128748	AC128748 Rattus no
	38	20	66.7	204302	2	AC068650	AC068650 Mus muscu
	39	20	66.7	216935	2	AC132648	AC132648 Rattus no
c	40	20	66.7	234787	2	AC073780	AC073780 Mus muscu
c	41	20	66.7	244520	2	AC128555	AC128555 Rattus no
	42	20	66.7	245754	2	AC127884	AC127884 Rattus no
	43	20	66.7	254067	2	AC133699	AC133699 Rattus no
	44	20	66.7	266396	2	AC108524	AC108524 Rattus no
c	45	20	66.7	315370	2	AC109390	AC109390 Rattus no

ALIGNMENTS

RESULT 1
E40089
LOCUS E40089 2048 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40089
VERSION E40089.1 GI:18627205
KEYWORDS JP 2000166577-A/3.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 2048)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 3 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000166577-A/3
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC C12N15/00,
PC C12N5/00,(C12N5/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1). .(2048).
FEATURES Location/Qualifiers
source 1. .2048
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 2048;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAATGTATGTCCGGTGTACATCTATGAC 30
||||||||||||||||||||||||||||||
Db 1742 AACAATGTATGTCCGGTGTACATCTATGAC 1771

```
RESULT 2
E40090
LOCUS       E40090                  2048 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40090
VERSION     E40090.1  GI:18627206
KEYWORDS    JP 2000166577-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2048)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 4 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/4
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2048).
FEATURES             Location/Qualifiers
     source          1. .2048
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 6;  Length 2048;
  Best Local Similarity   100.0%;  Pred. No. 0.00047;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||||||||||||||||||||||||||||
Db       1742 AACAATGTATGTCCGGTGTACATCTATGAC 1771


RESULT 3
E40087
LOCUS       E40087                  2052 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40087
VERSION     E40087.1  GI:18627203
KEYWORDS    JP 2000166577-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2052)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 1 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/1
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2052).
FEATURES             Location/Qualifiers
     source          1. .2052
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 6;  Length 2052;
  Best Local Similarity   100.0%;  Pred. No. 0.00047;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||||||||||||||||||||||||||||
Db       1746 AACAATGTATGTCCGGTGTACATCTATGAC 1775


RESULT 4
E40093
LOCUS       E40093                  2052 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40093
VERSION     E40093.1  GI:18627209
KEYWORDS    JP 2000166577-A/7.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2052)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 7 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
            PN   JP 2000166577-A/7
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2052).
FEATURES             Location/Qualifiers
     source          1. .2052
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 6;  Length 2052;
  Best Local Similarity   100.0%;  Pred. No. 0.00047;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||||||||||||||||||||||||||||
Db       1746 AACAATGTATGTCCGGTGTACATCTATGAC 1775


RESULT 5
E40091
LOCUS       E40091                  2056 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Plant promoter and terminator.
ACCESSION   E40091
VERSION     E40091.1  GI:18627207
KEYWORDS    JP 2000166577-A/5.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2056)
  AUTHORS   Nishikawa,S. and Oeda,K.
  TITLE     Plant promoter and terminator
  JOURNAL   Patent: JP 2000166577-A 5 20-JUN-2000;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L.
```

```
            PN   JP 2000166577-A/5
            PD   20-JUN-2000
            PF   01-OCT-1999 JP 1999281475
            PR
            PI   SATOMI NISHIKAWA,KENJI OEDA
            PC   C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
             C12N15/00,
            PC   C12N5/00,(C12N5/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   promoter        (1). .(2056).
FEATURES             Location/Qualifiers
     source          1. .2056
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 6;  Length 2056;
  Best Local Similarity   100.0%;  Pred. No. 0.00047;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||||||||||||||||||||||||||||
Db       1746 AACAATGTATGTCCGGTGTACATCTATGAC 1775


RESULT 6
AR076817
LOCUS       AR076817                2042 bp    DNA     linear   PAT 30-AUG-2000
DEFINITION  Sequence 2 from patent US 5959176.
ACCESSION   AR076817
VERSION     AR076817.1  GI:10003563
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2042)
  AUTHORS   Torikai,S. and Oeda,K.
  TITLE     Plant promoter and utilization thereof
  JOURNAL   Patent: US 5959176-A 2 28-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .2042
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             78.7%;  Score 23.6;  DB 6;  Length 2042;
  Best Local Similarity   86.7%;  Pred. No. 0.92;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              |||||  || ||||||||||| || ||||||
Db       1737 AACAACGTTTGTCCGGTGTATATTTATGAC 1766


RESULT 7
E15125
LOCUS       E15125                  2042 bp    DNA     linear   PAT 28-JUL-1999
DEFINITION  Promoter.
ACCESSION   E15125
VERSION     E15125.1  GI:5709808
KEYWORDS    JP 1998052273-A/2.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2042)
  AUTHORS   Torikai,T. and Oita,K.
  TITLE     VEGETABLE PROMOTER AND ITS USE
  JOURNAL   Patent: JP 1998052273-A 2 24-FEB-1998;
            SUMITOMO CHEM CO LTD
COMMENT     OS   Daucus carota L. (carrot)
            PN   JP 1998052273-A/2
            PD   24-FEB-1998
            PF   12-AUG-1996 JP 1996212680
            PI   TORIKAI TOSHIMI, OITA KENJI
            PC   C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
             strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2042
            FT                   /organism='Daucus carota L.'
            FT   promoter        1<. .<2042.
FEATURES             Location/Qualifiers
     source          1. .2042
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             78.7%;  Score 23.6;  DB 6;  Length 2042;
  Best Local Similarity   86.7%;  Pred. No. 0.92;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              |||||  || ||||||||||| || ||||||
Db       1737 AACAACGTTTGTCCGGTGTATATTTATGAC 1766


RESULT 8
AC125706
LOCUS       AC125706              348283 bp    DNA     linear   HTG 13-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-11J9, *** SEQUENCING IN PROGRESS ***,
            8 unordered pieces.
ACCESSION   AC125706
VERSION     AC125706.5  GI:24941053
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 348283)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
```

```
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 348283)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 348283)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 13, 2002 this sequence version replaced gi:23832296.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GDZZ
                Center clone name: CH230-11J9
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 272438 bases at least Q40
                Consensus quality: 277533 bases at least Q30
                Consensus quality: 280756 bases at least Q20
                Estimated insert size: 302017; sum-of-contigs estimation
                Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    11765: contig of 11765 bp in length
            *    11766    11865: gap of unknown length
            *    11866    48867: contig of 37002 bp in length
            *    48868    48967: gap of unknown length
            *    48968    54443: contig of 5476 bp in length
            *    54444    54543: gap of unknown length
            *    54544    59602: contig of 5059 bp in length
            *    59603    59702: gap of unknown length
            *    59703   343138: contig of 283436 bp in length
            *   343139   343238: gap of unknown length
            *   343239   344661: contig of 1423 bp in length
            *   344662   344761: gap of unknown length
            *   344762   345971: contig of 1210 bp in length
            *   345972   346071: gap of unknown length
            *   346072   348283: contig of 2212 bp in length.
FEATURES            Location/Qualifiers
     source         1. .348283
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-11J9"
     misc_feature   1. .1925
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   1909. .2527
                    /note="clone_boundary
                    clone_end:Sp6
                    site:
                    end_sequence:BH342645"
     misc_feature   8272. .11765
                    /note="wgs_contig"
     misc_feature   11866. .15419
                    /note="wgs_contig"
     misc_feature   20601. .22623
                    /note="wgs_contig"
     misc_feature   22674. .27148
                    /note="wgs_contig"
     misc_feature   48968. .52962
                    /note="wgs_contig"
     misc_feature   53013. .54443
                    /note="wgs_contig"
     misc_feature   54544. .56816
                    /note="wgs_contig"
     misc_feature   56870. .59602
                    /note="wgs_contig"
     misc_feature   59703. .61990
                    /note="wgs_contig"
     misc_feature   68117. .71973
                    /note="wgs_contig"
     misc_feature   338673. .340201
                    /note="wgs_contig"
     misc_feature   340622. .343138
                    /note="wgs_contig"
ORIGIN

  Query Match            72.7%;  Score 21.8;  DB 2;  Length 348283;
  Best Local Similarity   92.0%;  Pred. No. 12;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCT 25
              ||||||||||||| ||||||| |||
Db      90808 AACAATGTATGTCAGGTGTACGTCT 90832


RESULT 9
CR352244/c
LOCUS       CR352244              161762 bp    DNA     linear   HTG 27-MAR-2004
DEFINITION  Danio rerio clone CH211-131E11, *** SEQUENCING IN PROGRESS ***, 14
            unordered pieces.
ACCESSION   CR352244
VERSION     CR352244.4  GI:45772227
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
```

```
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 161762)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 27, 2004 this sequence version replaced gi:45598726.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zC131E11
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 157149 bases at least Q40
            Consensus quality: 157757 bases at least Q30
            Consensus quality: 158375 bases at least Q20
            Insert size: 160462; sum-of-contigs
            Insert size: 174990; 4.8% error; agarose-fp
            Quality coverage: 8.14x in Q20 bases; sum-of-contigs Quality
            coverage: 7.67x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    16722: contig of 16722 bp in length
            *    16723    16822: gap of 100 bp
            *    16823    28393: contig of 11571 bp in length
            *    28394    28493: gap of 100 bp
            *    28494    40368: contig of 11875 bp in length
            *    40369    40468: gap of 100 bp
            *    40469    57841: contig of 17373 bp in length
            *    57842    57941: gap of 100 bp
            *    57942    69071: contig of 11130 bp in length
            *    69072    69171: gap of 100 bp
            *    69172    77218: contig of 8047 bp in length
            *    77219    77318: gap of 100 bp
            *    77319    85619: contig of 8301 bp in length
            *    85620    85719: gap of 100 bp
            *    85720    89373: contig of 3654 bp in length
            *    89374    89473: gap of 100 bp
            *    89474    98585: contig of 9112 bp in length
            *    98586    98685: gap of 100 bp
            *    98686   103814: contig of 5129 bp in length
            *   103815   103914: gap of 100 bp
            *   103915   110517: contig of 6603 bp in length
            *   110518   110617: gap of 100 bp
            *   110618   139606: contig of 28989 bp in length
            *   139607   139706: gap of 100 bp
            *   139707   144556: contig of 4850 bp in length
            *   144557   144656: gap of 100 bp
            *   144657   161762: contig of 17106 bp in length.
FEATURES             Location/Qualifiers
     source          1. .161762
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-131E11"
                     /clone_lib="CHORI-211"
     misc_feature    1. .16722
                     /note="assembly_fragment:01696
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    16823. .28393
                     /note="assembly_fragment:00818
                     fragment_chain:1"
     misc_feature    28494. .40368
                     /note="assembly_fragment:00996
                     fragment_chain:1"
     misc_feature    40469. .57841
                     /note="assembly_fragment:01195
                     fragment_chain:1"
     misc_feature    57942. .69071
                     /note="assembly_fragment:00668
                     fragment_chain:1"
     misc_feature    69172. .77218
                     /note="assembly_fragment:00429
                     fragment_chain:1"
     misc_feature    77319. .85619
                     /note="assembly_fragment:00216
                     fragment_chain:1"
     misc_feature    85720. .89373
                     /note="assembly_fragment:00050
                     fragment_chain:1"
     misc_feature    89474. .98585
                     /note="assembly_fragment:00542
                     fragment_chain:1"
     misc_feature    98686. .103814
                     /note="assembly_fragment:00090
                     fragment_chain:1"
     misc_feature    103915. .110517
                     /note="assembly_fragment:00319
                     fragment_chain:1"
     misc_feature    110618. .139606
                     /note="assembly_fragment:02011
                     fragment_chain:2"
     misc_feature    139707. .144556
                     /note="assembly_fragment:00147
                     fragment_chain:2"
     misc_feature    144657. .161762
                     /note="assembly_fragment:01414
                     fragment_chain:2
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             70.7%;  Score 21.2;  DB 2;  Length 161762;
  Best Local Similarity   88.5%;  Pred. No. 23;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 AATGTATGTCCGGTGTACATCTATGA 29
              ||||| |||||||||| ||||| |||
Db       9417 AATGTGTGTCCGGTGTGCATCTGTGA 9392


RESULT 10
AC110515
LOCUS       AC110515               213033 bp    DNA     linear   HTG 09-MAR-2004
DEFINITION  Mus musculus chromosome 13 clone RP23-295C1 map 13, WORKING DRAFT
            SEQUENCE, 8 ordered pieces.
ACCESSION   AC110515
VERSION     AC110515.4  GI:44886734
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 213033)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 13, clone RP23-295C1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 213033)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
```

```
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 213033)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 2, 2004 this sequence version replaced gi:28460874.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L20363
                Center clone name: 295_C_1
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 210636 bases at least Q40
                Consensus quality: 211310 bases at least Q30
                Consensus quality: 211514 bases at least Q20
                Insert size: 186000; agarose-fp
                Insert size: 211645; sum-of-contigs
                Quality coverage: 11.5 in Q20 bases; agarose-fp
                Quality coverage: 10.1 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    10790: contig of 10790 bp in length
            *    10791    10890: gap of 100 bp
            *    10891    66637: contig of 55747 bp in length
            *    66638    66737: gap of 100 bp
            *    66738    69878: contig of 3141 bp in length
            *    69879    69978: gap of 100 bp
            *    69979    77514: contig of 7536 bp in length
            *    77515    77614: gap of 100 bp
            *    77615    88730: contig of 11116 bp in length
            *    88731    88830: gap of 100 bp
            *    88831    92930: contig of 4100 bp in length
            *    92931    93030: gap of 100 bp
            *    93031   188844: contig of 95814 bp in length
            *   188845   188944: gap of 100 bp
            *   188945   213033: contig of 24089 bp in length.
FEATURES             Location/Qualifiers
     source          1. .213033
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="13"
                     /map="13"
                     /clone="RP23-295C1"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1. .10790
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    10891. .66637
                     /note="assembly_fragment"
     misc_feature    66738. .69878
                     /note="assembly_fragment"
     misc_feature    69979. .77514
                     /note="assembly_fragment"
     misc_feature    77615. .88730
                     /note="assembly_fragment"
     misc_feature    88831. .92930
                     /note="assembly_fragment"
     misc_feature    93031. .188844
                     /note="assembly_fragment"
     misc_feature    188945. .213033
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             70.7%;  Score 21.2;  DB 2;  Length 213033;
  Best Local Similarity   88.5%;  Pred. No. 23;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 CAATGTATGTCCGGTGTACATCTATG 28
              | ||||||||||| ||||| ||||||
Db       9398 CTATGTATGTCCGATGTACGTCTATG 9423


RESULT 11
AC146937/c
LOCUS       AC146937              140142 bp    DNA     linear   PLN 19-JUN-2004
DEFINITION  Oryza sativa chromosome 11 BAC clone OSJNBa0004O15, complete
            sequence.
ACCESSION   AC146937
VERSION     AC146937.2  GI:48958697
KEYWORDS    HTG.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 140142)
  AUTHORS   Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Gansberger,K., Jones,K.M.,
            Overton II,L., Tsitrin,T., Kim,M., Bera,J., Jin,S., Fadrosh,D.W.,
            Tallon,L., Koo,H., Zismann,V., Hsiao,J., Blunt,S., Vanaken,S.,
            Riedmuller,S.B., Utterbach,T., Feldblyum,T., Yang,Q., Haas,B.,
            Suh,B., Peterson,J., Quackenbush,J., White,O., Salzberg,S. and
            Fraser,C.
  TITLE     Oryza sativa ssp. japonica cv. Nipponbare OSJNBa0004O15 BAC genomic
            sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 140142)
  AUTHORS   Buell,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-2003) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 140142)
  AUTHORS   Buell,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-OCT-2003) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   4  (bases 1 to 140142)
  AUTHORS   Buell,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2004) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     On Jun 19, 2004 this sequence version replaced gi:37991935.
FEATURES             Location/Qualifiers
     source          1. .140142
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="11"
                     /clone="OSJNBa0004O15"
                     /note="japonica cultivar-group"
ORIGIN

  Query Match             70.0%;  Score 21;  DB 8;  Length 140142;
  Best Local Similarity   82.8%;  Pred. No. 28;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||| | |||  ||||||||| ||||||
Db     113334 ACAATATTTGTTTGGTGTACATGTATGAC 113306


RESULT 12
AC025586
LOCUS       AC025586              185386 bp    DNA     linear   ROD 24-JUL-2002
DEFINITION  Genomic sequence for Mus musculus, clone RP23-320C8, complete
            sequence.
ACCESSION   AC025586
VERSION     AC025586.4  GI:21954008
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 185386)
  AUTHORS   McCombie,W.R., de la Bastide,M., Spiegel,L., Preston,R.,
            Nascimento,L., Zutavern,T., Balija,V., Bell,M., Miller,B.,
            Katzenberger,F., Muller,S., Sullivan,P., Yang,C., Dike,S.,
            Palmer,L., O'Shaughnessy,A. and Dedhia,N.
  TITLE     Genomic sequence for Mus musculus, clone RP23-320C8, complete
            sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 185386)
  AUTHORS   McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2000) Lita Annenberg Hazen Genome Sequencing
            Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
            Harbor, NY 11724, USA
REFERENCE   3  (bases 1 to 185386)
  AUTHORS   McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-2002) Lita Annenberg Hazen Genome Sequencing
            Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
            Harbor, NY 11724, USA
COMMENT     On Jul 24, 2002 this sequence version replaced gi:9972313.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest.
FEATURES             Location/Qualifiers
     source          1. .185386
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-320C8"
                     /clone_lib="RPCI-23"
     misc_feature    complement(23305. .23539)
                     /note="We believe the assembly to be correct.  The
                     sequence is a simple repeat (TTCCC)n in which the exact
                     number of repeat copies is unknown.  One high quality
                     subclone in the region spans the repeat into unique
                     sequence on both sides."
ORIGIN

  Query Match             70.0%;  Score 21;  DB 10;  Length 185386;
  Best Local Similarity   82.8%;  Pred. No. 29;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGA 29
              ||||||||| ||| ||| || ||||| ||
Db      98187 AACAATGTAAGTCTGGTTTAGATCTAAGA 98215


RESULT 13
AL592422
LOCUS       AL592422              205668 bp    DNA     linear   ROD 16-FEB-2002
DEFINITION  Mouse DNA sequence from clone RP23-218O16 on chromosome 11,
            complete sequence.
ACCESSION   AL592422
VERSION     AL592422.11  GI:18855205
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 205668)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 21, 2002 this sequence version replaced gi:17426517.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
```

```
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-218O16 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            This sequence is the entire insert of clone RP23-218O16.
FEATURES              Location/Qualifiers
     source           1. .205668
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
                      /chromosome="11"
                      /clone="RP23-218O16"
                      /clone_lib="RPCI-23"
     misc_feature     108825. .108843
                      /note="Single clone region. Assembly confirmed by
                      restriction digest data."
     misc_feature     108844
                      /note="Tandem repeat. Forced join. Gap size estimated to
                      be approximately 100bp by restriction digest data."
     misc_feature     108845. .108978
                      /note="Sequence from uni-directional dGTP big dye
                      terminator reads only"
     misc_feature     121988
                      /note="Tandem repeat. Forced join. Gap size estimated to
                      be approximately 200bp by restriction digest data."
ORIGIN

  Query Match             70.0%;  Score 21;  DB 10;  Length 205668;
  Best Local Similarity   82.8%;  Pred. No. 29;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGA 29
              ||||||||| ||| ||| || ||||| ||
Db     183167 AACAATGTAAGTCTGGTTTAGATCTAAGA 183195


RESULT 14
AL136130
LOCUS       AL136130               69674 bp    DNA     linear   PRI 04-MAR-2003
DEFINITION  Human DNA sequence from clone RP1-39M18 on chromosome 6q26-27,
            complete sequence.
ACCESSION   AL136130
VERSION     AL136130.7  GI:6911361
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 69674)
  AUTHORS   Laird,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 7, 2000 this sequence version replaced gi:6782223.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            ---------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ----------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            RP1-39M18 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2.
FEATURES              Location/Qualifiers
     source           1. .69674
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="6"
                      /map="q26-27"
                      /clone="RP1-39M18"
                      /clone_lib="RPCI-1"
ORIGIN

  Query Match             68.7%;  Score 20.6;  DB 9;  Length 69674;
  Best Local Similarity   85.2%;  Pred. No. 43;
  Matches   23;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              | || |||||||| |||||||| ||||
Db       8130 AAAACGTATGTCCAGTGTACATTTATG 8156


RESULT 15
AC023516
LOCUS       AC023516               78026 bp    DNA     linear   HTG 13-JUL-2000
DEFINITION  Homo sapiens clone RP11-115O4, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC023516
VERSION     AC023516.2  GI:9112723
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 78026)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-115O4
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 78026)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
```

```
            Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 13, 2000 this sequence version replaced gi:6978142.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L1909
                Center clone name: 115_O_4
            --------------.
            * NOTE: This record contains 84 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      800: contig of 800 bp in length
            *      801      900: gap of 100 bp
            *      901     1749: contig of 849 bp in length
            *     1750     1849: gap of 100 bp
            *     1850     2673: contig of 824 bp in length
            *     2674     2773: gap of 100 bp
            *     2774     3614: contig of 841 bp in length
            *     3615     3714: gap of 100 bp
            *     3715     4522: contig of 808 bp in length
            *     4523     4622: gap of 100 bp
            *     4623     5455: contig of 833 bp in length
            *     5456     5555: gap of 100 bp
            *     5556     6388: contig of 833 bp in length
            *     6389     6488: gap of 100 bp
            *     6489     7307: contig of 819 bp in length
            *     7308     7407: gap of 100 bp
            *     7408     8226: contig of 819 bp in length
            *     8227     8326: gap of 100 bp
            *     8327     9160: contig of 834 bp in length
            *     9161     9260: gap of 100 bp
            *     9261    10064: contig of 804 bp in length
            *    10065    10164: gap of 100 bp
            *    10165    10996: contig of 832 bp in length
            *    10997    11096: gap of 100 bp
            *    11097    11910: contig of 814 bp in length
            *    11911    12010: gap of 100 bp
            *    12011    12855: contig of 845 bp in length
            *    12856    12955: gap of 100 bp
            *    12956    13800: contig of 845 bp in length
            *    13801    13900: gap of 100 bp
            *    13901    14730: contig of 830 bp in length
            *    14731    14830: gap of 100 bp
            *    14831    15675: contig of 845 bp in length
            *    15676    15775: gap of 100 bp
            *    15776    16610: contig of 835 bp in length
            *    16611    16710: gap of 100 bp
            *    16711    17539: contig of 829 bp in length
            *    17540    17639: gap of 100 bp
            *    17640    18474: contig of 835 bp in length
            *    18475    18574: gap of 100 bp
            *    18575    19389: contig of 815 bp in length
            *    19390    19489: gap of 100 bp
            *    19490    20328: contig of 839 bp in length
            *    20329    20428: gap of 100 bp
            *    20429    21276: contig of 848 bp in length
            *    21277    21376: gap of 100 bp
            *    21377    22174: contig of 798 bp in length
            *    22175    22274: gap of 100 bp
            *    22275    23113: contig of 839 bp in length
            *    23114    23213: gap of 100 bp
            *    23214    24057: contig of 844 bp in length
            *    24058    24157: gap of 100 bp
            *    24158    25006: contig of 849 bp in length
            *    25007    25106: gap of 100 bp
            *    25107    25935: contig of 829 bp in length
            *    25936    26035: gap of 100 bp
            *    26036    26870: contig of 835 bp in length
            *    26871    26970: gap of 100 bp
            *    26971    27812: contig of 842 bp in length
            *    27813    27912: gap of 100 bp
            *    27913    28735: contig of 823 bp in length
            *    28736    28835: gap of 100 bp
            *    28836    29673: contig of 838 bp in length
            *    29674    29773: gap of 100 bp
            *    29774    30596: contig of 823 bp in length
            *    30597    30696: gap of 100 bp
            *    30697    31541: contig of 845 bp in length
            *    31542    31641: gap of 100 bp
            *    31642    32489: contig of 848 bp in length
            *    32490    32589: gap of 100 bp
            *    32590    33378: contig of 789 bp in length
            *    33379    33478: gap of 100 bp
            *    33479    34332: contig of 854 bp in length
            *    34333    34432: gap of 100 bp
            *    34433    35294: contig of 862 bp in length
            *    35295    35394: gap of 100 bp
            *    35395    36239: contig of 845 bp in length
            *    36240    36339: gap of 100 bp
            *    36340    37161: contig of 822 bp in length
            *    37162    37261: gap of 100 bp
            *    37262    38087: contig of 826 bp in length
            *    38088    38187: gap of 100 bp
            *    38188    39010: contig of 823 bp in length
            *    39011    39110: gap of 100 bp
            *    39111    39947: contig of 837 bp in length
            *    39948    40047: gap of 100 bp
            *    40048    40886: contig of 839 bp in length
            *    40887    40986: gap of 100 bp
            *    40987    41829: contig of 843 bp in length
            *    41830    41929: gap of 100 bp
            *    41930    42749: contig of 820 bp in length
            *    42750    42849: gap of 100 bp
            *    42850    43687: contig of 838 bp in length
            *    43688    43787: gap of 100 bp
            *    43788    44620: contig of 833 bp in length
            *    44621    44720: gap of 100 bp
            *    44721    45568: contig of 848 bp in length
            *    45569    45668: gap of 100 bp
            *    45669    46495: contig of 827 bp in length
            *    46496    46595: gap of 100 bp
            *    46596    47421: contig of 826 bp in length
            *    47422    47521: gap of 100 bp
            *    47522    48345: contig of 824 bp in length
            *    48346    48445: gap of 100 bp
            *    48446    49286: contig of 841 bp in length
            *    49287    49386: gap of 100 bp
            *    49387    50198: contig of 812 bp in length
            *    50199    50298: gap of 100 bp
            *    50299    51145: contig of 847 bp in length
            *    51146    51245: gap of 100 bp
            *    51246    52064: contig of 819 bp in length
            *    52065    52164: gap of 100 bp
            *    52165    53006: contig of 842 bp in length
            *    53007    53106: gap of 100 bp
            *    53107    53929: contig of 823 bp in length
            *    53930    54029: gap of 100 bp
```

```
            *    54030    54835: contig of 806 bp in length
            *    54836    54935: gap of 100 bp
            *    54936    55759: contig of 824 bp in length
            *    55760    55859: gap of 100 bp
            *    55860    56681: contig of 822 bp in length
            *    56682    56781: gap of 100 bp
            *    56782    57607: contig of 826 bp in length
            *    57608    57707: gap of 100 bp
            *    57708    58537: contig of 830 bp in length
            *    58538    58637: gap of 100 bp
            *    58638    59456: contig of 819 bp in length
            *    59457    59556: gap of 100 bp
            *    59557    60401: contig of 845 bp in length
            *    60402    60501: gap of 100 bp
            *    60502    61345: contig of 844 bp in length
            *    61346    61445: gap of 100 bp
            *    61446    62264: contig of 819 bp in length
            *    62265    62364: gap of 100 bp
            *    62365    63194: contig of 830 bp in length
            *    63195    63294: gap of 100 bp
            *    63295    64128: contig of 834 bp in length
            *    64129    64228: gap of 100 bp

  Query Match             68.7%;  Score 20.6;  DB 2;  Length 78026;
  Best Local Similarity   85.2%;  Pred. No. 43;
  Matches   23;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              | || |||||||| |||||||| ||||
Db      40150 AAAACGTATGTCCAGTGTACATTTATG 40176


Search completed: December  6, 2004, 22:47:35
Job time : 142.012 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 15:16:44 ; Search time 16.5905 Seconds
(without alignments)
9492.325 Million cell updates/sec

Title: US-09-806-197-24
Perfect score: 30
Sequence: 1 aacaatgtatgtccggtgtacatctatgac 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30	100.0	30	3	AAA37981	Aaa37981 PCR prime
	2	30	100.0	2048	3	AAA37962	Aaa37962 Plasmid #
	3	30	100.0	2048	3	AAA37963	Aaa37963 Plasmid #
	4	30	100.0	2052	3	AAA37959	Aaa37959 Carrot pr
	5	30	100.0	2052	3	AAA37961	Aaa37961 Carrot pr
	6	30	100.0	2056	3	AAA37964	Aaa37964 Plasmid #
	7	23.6	78.7	2042	2	AAV15144	Aav15144 New promo
c	8	19.6	65.3	2931	12	ADN73150	Adn73150 Thale cre
c	9	19	63.3	2700	10	ADB53615	Adb53615 Primary r
	10	19	63.3	35042	3	AAA81454	Aaa81454 N. mening
	11	19	63.3	110000	3	AAA81489_1	Continuation (2 of
	12	19	63.3	236303	4	AAS11614	Aas11614 Human gen
	13	19	63.3	349980	3	AAF21611	Aaf21611 Neisseria
c	14	18.8	62.7	1630	6	AAS62657	Aas62657 cDNA sequ
	15	18.6	62.0	1569	3	AAA30829	Aaa30829 Zebrafish
	16	18.6	62.0	1671	3	AAA48445	Aaa48445 Zebrafish
	17	18.6	62.0	1671	10	ADC42329	Adc42329 cDNA enco
	18	18.6	62.0	2152	3	AAA48446	Aaa48446 Zebrafish
	19	18.6	62.0	2152	10	ADC42306	Adc42306 cDNA enco
	20	18.6	62.0	2152	10	ADH61060	Adh61060 Zebrafish
	21	18.6	62.0	90442	9	ADA03077	Ada03077 Mouse mCG
	22	18.6	62.0	90442	9	ADA66361	Ada66361 Mouse mCG
	23	18.6	62.0	90442	10	ADB72815	Adb72815 Mouse mCG
	24	18.6	62.0	90442	10	ADC26997	Adc26997 Mouse car
	25	18.6	62.0	90442	11	ADL27155	Adl27155 Mouse gen
	26	18.4	61.3	3177	4	ABL28637	Abl28637 Drosophil
	27	18.4	61.3	4354	4	ABL11038	Abl11038 Drosophil
	28	18.4	61.3	5460	4	ABL28636	Abl28636 Drosophil
	29	18.2	60.7	4919	4	ABL15204	Abl15204 Drosophil
c	30	18.2	60.7	40050	9	ADA02585	Ada02585 Mouse Ics
c	31	18.2	60.7	40050	10	ADB72323	Adb72323 Mouse Ics
c	32	18.2	60.7	40050	10	ADE95833	Ade95833 Mouse ics
	33	18	60.0	287	12	ADL11418	Adl11418 Cat flea
	34	18	60.0	287	12	ADL11392	Adl11392 Cat flea
	35	18	60.0	386	12	ADL11486	Adl11486 Cat flea
c	36	18	60.0	401	4	AAK96285	Aak96285 Human neu
c	37	18	60.0	401	4	AAK97778	Aak97778 Human neu
c	38	18	60.0	401	6	ABT01055	Abt01055 Human neu
c	39	18	60.0	401	6	ABT02548	Abt02548 Human neu
	40	18	60.0	448	12	ADL10839	Adl10839 Cat flea
c	41	18	60.0	83709	12	ADQ19964	Adq19964 Human sof
c	42	18	60.0	110000	4	AAK95240_10	Continuation (11 o
c	43	18	60.0	110000	4	AAK95240_11	Continuation (12 o
c	44	18	60.0	110000	4	AAK96733_10	Continuation (11 o
c	45	18	60.0	110000	4	AAK96733_11	Continuation (12 o

ALIGNMENTS

RESULT 1
AAA37981
ID AAA37981 standard; DNA; 30 BP.
XX
AC AAA37981;
XX
DT 18-AUG-2000 (first entry)
XX
DE PCR primer S used for carrot promoter amplification.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility;
KW PCR primer; ss.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 6; Page 44; 81pp; English.
XX
CC This sequence represents a PCR primer used to amplify a carrot promoter
CC sequence. The invention relates to plant promoters and terminators from
CC Daucus carota L. which are capable of expressing a gene of interest in
CC plants. The invention also includes a chimeric gene characterized in that
CC it comprises the promoter and a desired gene linked to each other in the
CC form capable of functioning. A method of producing a transformant
CC comprises introducing the promoter, the chimeric gene or a vector
CC comprising the promoter and a desired gene or terminator sequence into a
CC host cell. The plant promoters and terminators are useful in plant
CC breeding, for e.g. fertilities of plants may be controlled by expressing,

```
CC   in the host cells, a sense or antisense gene of a male sterility related
CC   gene such as S-locus-specific RNase gene
XX
SQ   Sequence 30 BP; 9 A; 6 C; 6 G; 9 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 3;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 0.00016;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||||||||||||||||||||||||||||
Db          1 AACAATGTATGTCCGGTGTACATCTATGAC 30


RESULT 2
AAA37962
ID   AAA37962 standard; DNA; 2048 BP.
XX
AC   AAA37962;
XX
DT   18-AUG-2000  (first entry)
XX
DE   Plasmid #1 DNA sequence used in mutation of promoter sequence.
XX
KW   Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS   Daucus carota.
XX
PN   WO200020613-A1.
XX
PD   13-APR-2000.
XX
PF   28-SEP-1999;   99WO-JP005303.
XX
PR   02-OCT-1998;   98JP-00281124.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Nishikawa S,  Oeda K;
XX
DR   WPI; 2000-303791/26.
XX
PT   New Plant promoters and terminators from Daucus carota L., useful in
PT   plant breeding, for e.g. controlling fertilities of plants.
XX
PS   Example 8; Page 71-73; 81pp; English.
XX
CC   This sequence represents a plasmid sequence used in a method for
CC   introducing a mutation into a carrot promoter. The invention relates to
CC   plant promoters and terminators from Daucus carota L. which are capable
CC   of expressing a gene of interest in plants. The invention also includes a
CC   chimeric gene characterized in that it comprises the promoter and a
CC   desired gene linked to each other in the form capable of functioning. A
CC   method of producing a transformant comprises introducing the promoter,
CC   the chimeric gene or a vector comprising the promoter and a desired gene
CC   or terminator sequence into a host cell. The plant promoters and
CC   terminators are useful in plant breeding, for e.g. fertilities of plants
CC   may be controlled by expressing, in the host cells, a sense or antisense
CC   gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ   Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 3;  Length 2048;
  Best Local Similarity   100.0%;  Pred. No. 0.00036;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||||||||||||||||||||||||||||
Db       1742 AACAATGTATGTCCGGTGTACATCTATGAC 1771


RESULT 3
AAA37963
ID   AAA37963 standard; DNA; 2048 BP.
XX
AC   AAA37963;
XX
DT   18-AUG-2000  (first entry)
XX
DE   Plasmid #2 DNA sequence used in mutation of promoter sequence.
XX
KW   Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS   Daucus carota.
XX
PN   WO200020613-A1.
XX
PD   13-APR-2000.
XX
PF   28-SEP-1999;   99WO-JP005303.
XX
PR   02-OCT-1998;   98JP-00281124.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Nishikawa S,  Oeda K;
XX
DR   WPI; 2000-303791/26.
XX
PT   New Plant promoters and terminators from Daucus carota L., useful in
PT   plant breeding, for e.g. controlling fertilities of plants.
XX
PS   Example 8; Page 73-74; 81pp; English.
XX
CC   This sequence represents a plasmid sequence used in a method for
CC   introducing a mutation into a carrot promoter. The invention relates to
CC   plant promoters and terminators from Daucus carota L. which are capable
CC   of expressing a gene of interest in plants. The invention also includes a
CC   chimeric gene characterized in that it comprises the promoter and a
CC   desired gene linked to each other in the form capable of functioning. A
CC   method of producing a transformant comprises introducing the promoter,
CC   the chimeric gene or a vector comprising the promoter and a desired gene
CC   or terminator sequence into a host cell. The plant promoters and
CC   terminators are useful in plant breeding, for e.g. fertilities of plants
CC   may be controlled by expressing, in the host cells, a sense or antisense
CC   gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ   Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 3;  Length 2048;
  Best Local Similarity   100.0%;  Pred. No. 0.00036;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||||||||||||||||||||||||||||
Db       1742 AACAATGTATGTCCGGTGTACATCTATGAC 1771


RESULT 4
AAA37959
ID   AAA37959 standard; DNA; 2052 BP.
XX
AC   AAA37959;
XX
DT   18-AUG-2000  (first entry)
XX
DE   Carrot promoter sequence #1.
XX
KW   Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS   Daucus carota.
XX
PN   WO200020613-A1.
XX
PD   13-APR-2000.
XX
```

PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 69-70; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 3; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAATGTATGTCCGGTGTACATCTATGAC 30
||||||||||||||||||||||||||||||
Db 1746 AACAATGTATGTCCGGTGTACATCTATGAC 1775

RESULT 5
AAA37961
ID AAA37961 standard; DNA; 2052 BP.
XX
AC AAA37961;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #2.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 78-79; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 736 A; 318 C; 317 G; 681 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 3; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAATGTATGTCCGGTGTACATCTATGAC 30
||||||||||||||||||||||||||||||
Db 1746 AACAATGTATGTCCGGTGTACATCTATGAC 1775

RESULT 6
AAA37964
ID AAA37964 standard; DNA; 2056 BP.
XX
AC AAA37964;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #3 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 74-76; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 3; Length 2056;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||||||||||||||||||||||||||||
Db       1746 AACAATGTATGTCCGGTGTACATCTATGAC 1775
```

RESULT 7
AAV15144
ID AAV15144 standard; DNA; 2042 BP.
XX
AC AAV15144;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX
PN EP824150-A2.
XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Torikai S, Oeda K;
XX
DR WPI; 1998-122310/12.
XX
PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
PS Claim 2; Page 15-16; 31pp; English.
XX
CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;

```
  Query Match             78.7%;  Score 23.6;  DB 2;  Length 2042;
  Best Local Similarity   86.7%;  Pred. No. 0.44;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||| || |||||||||| || ||||||
Db       1737 AACAACGTTTGTCCGGTGTATATTTATGAC 1766
```

RESULT 8
ADN73150/c
ID ADN73150 standard; cDNA; 2931 BP.
XX
AC ADN73150;
XX
DT 15-JUL-2004 (first entry)
XX
DE Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 1045.
XX
KW gene; ss; plant; transgenic; E2Fa/DPa transcription factor;
KW growth regulator; animal feed product; thale cress;
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX
OS Arabidopsis thaliana.
XX
PN WO2004035798-A2.
XX
PD 29-APR-2004.
XX
PF 20-OCT-2003; 2003WO-EP011658.
XX
PR 18-OCT-2002; 2002EP-00079408.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Inze D, De Veylder L, Vlieghe K;
XX
DR WPI; 2004-348466/32.
DR P-PSDB; ADN73151.
XX
PT Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.
XX
PS Claim 1; SEQ ID NO 1045; 134pp; English.
XX
CC This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/DPa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/ or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress cDNA
CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/DPa
CC transcription factor, given in an exemplification of the invention.
XX
SQ Sequence 2931 BP; 818 A; 547 C; 682 G; 884 T; 0 U; 0 Other;

```
  Query Match             65.3%;  Score 19.6;  DB 12;  Length 2931;
  Best Local Similarity   84.6%;  Pred. No. 40;
  Matches   22;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 ATGTATGTCCGGTGTACATCTATGAC 30
              | ||| |||  ||||||||||||||||
Db       2591 AAGTACGTCTCGTGTACATCTATGAC 2566
```

RESULT 9
ADB53615/c
ID ADB53615 standard; DNA; 2700 BP.
XX
AC ADB53615;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4157.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX

PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 4157; 874pp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 2700 BP; 791 A; 633 C; 691 G; 585 T; 0 U; 0 Other;

Query Match 63.3%; Score 19; DB 10; Length 2700;
Best Local Similarity 81.5%; Pred. No. 77;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 AATGTATGTCCGGTGTACATCTATGAC 30
||||| |||||| || |||||| | ||
Db 1637 AATGTCTGTCCGCTGCACATCTCTCAC 1611

RESULT 10
AAA81454
ID AAA81454 standard; DNA; 35042 BP.
XX
AC AAA81454;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_2 SEQ ID NO:2.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-0103794P.
PR 30-APR-1999; 99US-0132068P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX
PS Claim 7; Page 242-253; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neissariae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 35042 BP; 8742 A; 9703 C; 8287 G; 8310 T; 0 U; 0 Other;

Query Match 63.3%; Score 19; DB 3; Length 35042;
Best Local Similarity 81.5%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AACAATGTATGTCCGGTGTACATCTAT 27
||||||||||| || ||| |||||||
Db 8712 AACAATGTATTTCTATTGTTCATCTAT 8738

RESULT 11
AAA81489_1
Continuation (2 of 9) of AAA81489 from base 100001 (N. meningitidis partial DNA sequenc
WP Sequence split into 9 fragments LOCUS AAA81489 Accession Aaa81489
WP Fragment Name Begin End
WP AAA81489_0 1 110000
WP AAA81489_1 100001 210000

```
WP      AAA81489_2          200001    310000
WP      AAA81489_3          300001    410000
WP      AAA81489_4          400001    510000
WP      AAA81489_5          500001    610000
WP      AAA81489_6          600001    710000
WP      AAA81489_7          700001    810000
WP      AAA81489_8          800001    837096

  Query Match             63.3%;  Score 19;  DB 3;  Length 110000;
  Best Local Similarity   81.5%;  Pred. No. 1.6e+02;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTAT 27
              |||||||||| ||   ||| |||||||
Db      16813 AACAATGTATTTCTATTGTTCATCTAT 16839


RESULT 12
AAS11614
ID   AAS11614 standard; DNA; 236303 BP.
XX
AC   AAS11614;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human genomic DNA containing exons 2-17 of the CRIM1 gene.
XX
KW   CRIM-1; Human; human chromosome 2p21-16.3; ophthalmological;
KW   neuroprotective; renal; osteopathic; dental; vulnerary; immunogen;
KW   antibody; gene therapy; neurodegenerative disease; eye disorder;
KW   cataract; bone morphogenic protein; BMG; renal disease; bone abnormality;
KW   tooth abnormality; wound; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            33104. .33277
FT                   /*tag= a
FT                   /number= 2
FT   intron          33278. .77746
FT                   /*tag= b
FT                   /number= 2
FT   exon            77747. .77989
FT                   /*tag= c
FT                   /number= 3
FT   intron          77990. .79103
FT                   /*tag= d
FT                   /number= 3
FT   exon            79104. .79224
FT                   /*tag= e
FT                   /number= 4
FT   intron          79225. .101022
FT                   /*tag= f
FT                   /number= 4
FT   exon            101023. .101144
FT                   /*tag= g
FT                   /number= 5
FT   intron          101145. .113377
FT                   /*tag= h
FT                   /number= 5
FT   exon            113378. .113560
FT                   /*tag= i
FT                   /number= 6
FT   intron          113561. .115985
FT                   /*tag= j
FT                   /number= 6
FT   exon            115986. .116183
FT                   /*tag= k
FT                   /number= 7
FT   intron          116184. .135707
FT                   /*tag= l
FT                   /number= 7
FT   exon            135708. .135836
FT                   /*tag= m
FT                   /number= 8
FT   intron          135837. .146471
FT                   /*tag= m
FT                   /number= 8
FT   exon            146472. .146628
FT                   /*tag= o
FT                   /number= 9
FT   intron          146629. .148761
FT                   /*tag= p
FT                   /number= 9
FT   exon            148762. .148883
FT                   /*tag= q
FT                   /number= 10
FT   intron          148884. .150044
FT                   /*tag= r
FT                   /number= 10
FT   exon            150045. .150254
FT                   /*tag= s
FT                   /number= 11
FT   intron          150255. .153815
FT                   /*tag= t
FT                   /number= 11
FT   exon            153816. .154031
FT                   /*tag= u
FT                   /number= 12
FT   intron          154032. .158580
FT                   /*tag= v
FT                   /number= 12
FT   exon            158581. .158802
FT                   /*tag= w
FT                   /number= 13
FT   intron          158803. .173982
FT                   /*tag= x
FT                   /number= 13
FT   exon            173983. .174177
FT                   /*tag= y
FT                   /number= 14
FT   intron          174178. .181006
FT                   /*tag= z
FT                   /number= 14
FT   exon            181007. .181129
FT                   /*tag= aa
FT                   /number= 15
FT   intron          181130. .183612
FT                   /*tag= ab
FT                   /number= 15
FT   exon            183613. .183800
FT                   /*tag= ac
FT                   /number= 16
FT   intron          183801. .185152
FT                   /*tag= ad
FT                   /number= 16
FT   exon            185153. .187765
FT                   /*tag= ae
FT                   /number= 17
XX
PN   WO200138519-A1.
XX
PD   31-MAY-2001.
XX
PF   24-NOV-2000; 2000WO-AU001435.
XX
PR   26-NOV-1999;   99AU-00004348.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Little M,  Yamada T,  Holmes G,  Georgas K,  Kolle G,  Wilkinson L;
XX
DR   WPI; 2001-343951/36.
XX
PT   Nucleic acids from human chromosome 2p21-16.3 and the encoded peptide,
PT   useful for preventing, diagnosing and treating e.g. eye disease,
```

PT especially cataract formation.
XX
PS Claim 4; Fig 3; 169pp; English.
XX
CC The invention relates to nucleic acids from human chromosome 2p21-16.3
CC and the encoded peptide (and mouse and chicken orthologues) that
CC comprises a PGECCPLP group, an insulin-like growth factor binding protein
CC (IGFBP)-like domain, cysteine-rich domains, an RGD (undefined) group and
CC a transmembrane domain. The protein, e.g. CRIM1, interacts with peptides
CC of the transforming growth factor superfamily. A composition comprising
CC an expression construct comprising the nucleic acids of the invention or
CC a mimetic which antagonises or mimics an activity of a CRIM1 polypeptide
CC may be used in a method for modulating the biological activity of a
CC polypeptide of the bone morphogenic protein (BMP) family. In this way
CC they may be used to prevent or treat an eye disease, especially cataract
CC formation. They may also be used to treat neurodegenerative diseases,
CC renal and kidney disease, bone and tooth abnormalities, wounds and skin
CC damage, e.g. by use of the nucleic acid in gene therapy by using
CC antibodies directed against CRIM1 polypeptides. The present sequence is a
CC Human genomic DNA containing exons 2-17 of the CRIM1 gene
XX
SQ Sequence 236303 BP; 70199 A; 46544 C; 47996 G; 71563 T; 0 U; 1 Other;

Query Match 63.3%; Score 19; DB 4; Length 236303;
Best Local Similarity 81.5%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 AATGTATGTCCGGTGTACATCTATGAC 30
||||| ||| | |||||||||| ||||
Db 133196 AATGTGTGTTAGATGTACATCTGTGAC 133222

RESULT 13
AAF21611
ID AAF21611 standard; DNA; 349980 BP.
XX
AC AAF21611;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US005928.
XX
PR 30-APR-1999; 99US-0132068P.
PR 08-OCT-1999; 99WO-US023573.
PR 28-FEB-2000; 2000GB-00004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections.
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of Neisseria
CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
CC represent fragments of the NMB genomic sequence, as the sequence was too
CC long to go in a record on its own it was split into 8 sequences which
CC overlap each other at the beginning and end of each sequence by 49980 bp
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
CC primers which are used in the exemplification of the present invention.
CC The NMB genome and fragments from it have antibacterial activity, and can
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
CC and/or antibodies which binds to the proteins can be used in compositions
CC for treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find antigenic
CC or immunogenic proteins which are more effective in vaccines than the
CC outer membrane proteins currently used
XX
SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;

Query Match 63.3%; Score 19; DB 3; Length 349980;
Best Local Similarity 81.5%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AACAATGTATGTCCGGTGTACATCTAT 27
|||||||||| || ||| |||||||
Db 52042 AACAATGTATTTCTATTGTTCATCTAT 52068

RESULT 14
AAS62657/c
ID AAS62657 standard; cDNA; 1630 BP.
XX
AC AAS62657;
XX
DT 14-FEB-2002 (first entry)
XX
DE cDNA sequence #444 encoding novel human secreted protein.
XX
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW immunosuppressive; antirheumatic; ss.
XX
OS Homo sapiens.
XX
PN WO200177291-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US010485.
XX
PR 06-APR-2000; 2000US-0195604P.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
XX
DR WPI; 2002-010900/01.
XX
PT New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease.
XX
PS Claim 1; Page 308; 391pp; English.
XX
CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides a
CC method for producing proteins from these polynucleotide sequences. The

```
CC   proteins are useful for identifying compounds that modulate their
CC   activity and production, and the cell is also useful for identifying
CC   compounds that modulate expression of the polynucleotide sequences
CC   encoding the secreted proteins. The sequences of the invention are useful
CC   for treating diseases such as hyperproliferative disorders (e.g. cancer),
CC   immune deficiency disorders (e.g. severe combined immunodeficiency
CC   (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
CC   (e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis) and
CC   infectious disorders (e.g. hepatitis). The polynucleotide sequences of
CC   the invention are also useful in gene therapy. AAS62214-AAS62838
CC   represent the cDNA sequences of the invention that encode for novel human
CC   secreted proteins
XX
SQ   Sequence 1630 BP; 409 A; 346 C; 284 G; 590 T; 0 U; 1 Other;

  Query Match             62.7%;  Score 18.8;  DB 6;  Length 1630;
  Best Local Similarity   76.7%;  Pred. No. 87;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||| | ||| | || ||| ||||||||||
Db        633 AACGAAGTAAGCCCTGTGATCATCTATGAC 604


RESULT 15
AAA30829
ID   AAA30829 standard; DNA; 1569 BP.
XX
AC   AAA30829;
XX
DT   15-SEP-2003  (revised)
DT   29-AUG-2000  (first entry)
XX
DE   Zebrafish PTH3R receptor coding sequence.
XX
KW   Zebrafish; PTH1R receptor; PTH3R receptor; diagnosis; cancer;
KW   parathyroid hormone type 1 receptor; parathyroid hormone type 3 receptor;
KW   ds.
XX
OS   Danio rerio.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1569
FT                   /*tag= a
FT                   /product= "PTH3R receptor"
FT                   /partial
FT                   /note= "no stop codon given"
XX
PN   WO200032771-A1.
XX
PD   08-JUN-2000.
XX
PF   28-MAY-1999;   99WO-US011883.
XX
PR   30-NOV-1998;   98US-0110467P.
XX
PA   (JUEP/) JUEPPNER H.
PA   (RUBI/) RUBIN D A.
XX
PI   Jueppner H,  Rubin DA;
XX
DR   WPI; 2000-412319/35.
DR   P-PSDB; AAY90231.
XX
PT   Novel zebrafish parathyroid hormone/parathyroid hormone related peptide
PT   receptor 3 and isolated nucleic acid encoding zebrafish parathyroid
PT   hormone receptor 1 for treating disorders associated with receptor
PT   function.
XX
PS   Claim 23; Fig 1d; 111pp; English.
XX
CC   This sequence encodes a parathyroid hormone receptor type 3 (PTH3R)
CC   receptor protein of the invention. The invention also relates to a PTH1R
CC   receptor protein. Antagonists of PTH1R or PTH3R can be used for the
CC   treatment of diseases associated with an increase in PTH1R or PTH3R
CC   activity, respectively. The peptides are used for diagnosis or prognosis
CC   of diseases and disorders associated with PTH3R or PTH1R, such as cancer.
CC   The polypeptides can be used as a molecular weight markers on sodium
CC   dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels, or
CC   on molecular sieve gel filtration columns. Antigenic epitope-bearing
CC   peptides and polypeptides are useful to raise antibodies, including
CC   monoclonal antibodies, that bind specifically to a polypeptide. The
CC   peptides are useful during diagnosis of diseases and disorders in mammals
CC   involving PTH1R or PTH3R receptor expression or function. Mutations that
CC   affect PTH1R or PTH3R sequence and/or expression levels of PTH1R or PTH3R
CC   could be diagnostic for patients with disease or disorders of a
CC   developmental, physiological or neurological nature. The nucleic acid
CC   molecules are valuable for chromosome identification. The mapping of DNAs
CC   to chromosomes is an important first step in correlating those sequences
CC   with genes associated with disease. (Updated on 15-SEP-2003 to
CC   standardise OS field)
XX
SQ   Sequence 1569 BP; 382 A; 358 C; 418 G; 411 T; 0 U; 0 Other;

  Query Match             62.0%;  Score 18.6;  DB 3;  Length 1569;
  Best Local Similarity   84.0%;  Pred. No. 1.1e+02;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 TGTATGTCCGGTGTACATCTATGAC 30
              | | ||||| | ||||||||||||
Db        246 TCTGTGTCCTGAGTACATCTATGAC 270


Search completed: December  6, 2004, 19:02:31
Job time : 19.5905 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:08:52 ; Search time 3.00716 Seconds
(without alignments)
7090.965 Million cell updates/sec

Title: US-09-806-197-24
Perfect score: 30
Sequence: 1 aacaatgtatgtccggtgtacatctatgac 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	23.6	78.7	2042	2	US-08-911-434A-2	Sequence 2, Appli
	2	18.6	62.0	2152	4	US-09-449-632-3	Sequence 3, Appli
	3	17.8	59.3	438	3	US-08-651-155B-134	Sequence 134, App
	4	17.8	59.3	438	4	US-09-194-036B-134	Sequence 134, App
	5	17.8	59.3	908	3	US-09-457-046B-15	Sequence 15, Appl
c	6	17.4	58.0	198	4	US-09-248-796A-7260	Sequence 7260, Ap
c	7	17.4	58.0	1507	3	US-08-605-150A-15	Sequence 15, Appl
c	8	17.4	58.0	2182	4	US-09-221-017B-916	Sequence 916, App
c	9	17.4	58.0	2662	3	US-08-750-357-8	Sequence 8, Appli
c	10	17.4	58.0	3848	3	US-09-112-096-28	Sequence 28, Appl
c	11	17.4	58.0	5668	3	US-09-112-096-14	Sequence 14, Appl
c	12	17.4	58.0	5668	4	US-09-636-215-777	Sequence 777, App
c	13	17.4	58.0	5668	4	US-09-685-166A-777	Sequence 777, App
c	14	17.4	58.0	5668	4	US-09-679-426-777	Sequence 777, App
	15	17.4	58.0	48974	3	US-08-920-422-17	Sequence 17, Appl
c	16	17.2	57.3	351	4	US-09-328-352-182	Sequence 182, App
	17	17.2	57.3	5474	1	US-08-137-252-2	Sequence 2, Appli
	18	17.2	57.3	11663	1	US-08-446-932-1	Sequence 1, Appli
	19	17.2	57.3	11663	1	US-08-801-263A-1	Sequence 1, Appli
	20	17.2	57.3	11663	1	US-08-801-263A-7	Sequence 7, Appli
	21	17.2	57.3	11663	3	US-09-102-248-1	Sequence 1, Appli
	22	17.2	57.3	11663	3	US-09-102-248-7	Sequence 7, Appli
	23	17.2	57.3	11663	4	US-09-367-764-1	Sequence 1, Appli
	24	17.2	57.3	11663	4	US-09-367-764-7	Sequence 7, Appli
	25	17	56.7	745	4	US-09-270-767-12913	Sequence 12913, A
	26	17	56.7	11580	3	US-09-334-220-4	Sequence 4, Appli
	27	16.8	56.0	275	1	US-07-789-919A-1	Sequence 1, Appli
	28	16.8	56.0	275	1	US-08-209-846A-1	Sequence 1, Appli
	29	16.8	56.0	275	2	US-08-472-809B-1	Sequence 1, Appli
	30	16.8	56.0	275	3	US-08-438-265-1	Sequence 1, Appli
c	31	16.8	56.0	1529	3	US-08-821-278A-1	Sequence 1, Appli
	32	16.8	56.0	1609	4	US-09-449-632-1	Sequence 1, Appli
c	33	16.8	56.0	2166	4	US-09-107-532A-3235	Sequence 3235, Ap
c	34	16.8	56.0	4819	4	US-09-774-528-72	Sequence 72, Appl
	35	16.8	56.0	5707	2	US-08-472-809B-8	Sequence 8, Appli
	36	16.8	56.0	6345	2	US-08-472-809B-7	Sequence 7, Appli
c	37	16.8	56.0	640681	4	US-09-790-988-1	Sequence 1, Appli
	38	16.4	54.7	148	4	US-09-513-999C-30057	Sequence 30057, A
c	39	16.4	54.7	484	3	US-08-945-983-3	Sequence 3, Appli
	40	16.4	54.7	484	3	US-08-945-983-5	Sequence 5, Appli
	41	16.4	54.7	600	6	5198345-4	Patent No. 5198345
c	42	16.4	54.7	610	4	US-09-270-767-6688	Sequence 6688, Ap
c	43	16.4	54.7	610	4	US-09-270-767-21970	Sequence 21970, A
c	44	16.4	54.7	3349	4	US-09-375-318-36	Sequence 36, Appl
	45	16.4	54.7	51259	3	US-08-781-891-209	Sequence 209, App

ALIGNMENTS

```
RESULT 1
US-08-911-434A-2
; Sequence 2, Application US/08911434A
; Patent No. 5959176
;  GENERAL INFORMATION:
;    APPLICANT:  TORIKAI, Satomi
;    APPLICANT:  OEDA, Kenji
;    TITLE OF INVENTION:  PLANT PROMOTER AND UTILIZATION THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH, LLP
;      STREET:  P.O. BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VIRGINIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  22040
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/911,434A
;      FILING DATE:  12-AUG-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Raymond C.
;      REGISTRATION NUMBER:  21,066
;      REFERENCE/DOCKET NUMBER:  2185-0199P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)205-8000
;      TELEFAX:  (703)205-8050
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2042 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Daucus carota L.
;      INDIVIDUAL ISOLATE:  Kuroda Gosun
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  1..2042
US-08-911-434A-2

  Query Match            78.7%;  Score 23.6;  DB 2;  Length 2042;
  Best Local Similarity  86.7%;  Pred. No. 0.12;
```

```
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||||| || |||||||||||| || ||||||
Db       1737 AACAACGTTTGTCCGGTGTATATTTATGAC 1766


RESULT 2
US-09-449-632-3
; Sequence 3, Application US/09449632
; Patent No. 6541220
; GENERAL INFORMATION:
;  APPLICANT: Jppner, Harald
;  APPLICANT:  Rubin, David A.
;  TITLE OF INVENTION: PTH1R and PTH3R Receptors, Methods and Uses Thereof
;  FILE REFERENCE: 0609.4740001/SRL/M-G
;  CURRENT APPLICATION NUMBER: US/09/449,632
;  CURRENT FILING DATE:  1999-11-30
;  PRIOR APPLICATION NUMBER: US 60/110,467
;  PRIOR FILING DATE: 1998-11-30
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;    LENGTH: 2152
;    TYPE: DNA
;    ORGANISM: zebrafish
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (394)..(2019)
;    NAME/KEY: misc_feature
;    LOCATION: (2125)..(2125)
;    OTHER INFORMATION: n is any nucleotide of a,t,g or c
US-09-449-632-3

  Query Match             62.0%;  Score 18.6;  DB 4;  Length 2152;
  Best Local Similarity   84.0%;  Pred. No. 24;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 TGTATGTCCGGTGTACATCTATGAC 30
              | | ||||| | ||||||||||||||
Db        639 TCTGTGTCCTGAGTACATCTATGAC 663


RESULT 3
US-08-651-155B-134
; Sequence 134, Application US/08651155B
; Patent No. 6365401
;  GENERAL INFORMATION:
;    APPLICANT:  Mahan Dr., Michael J.
;    APPLICANT:  Conner Mr., Christopher P.
;    APPLICANT:  Hiethoff Mr., Douglas M.
;    TITLE OF INVENTION:  METHOD AND PROBES FOR THE IDENTIFICATION
;    TITLE OF INVENTION:  OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST
;    TITLE OF INVENTION:  INFECTION
;    NUMBER OF SEQUENCES:  255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chrisman, Bynum & Johnson, P.C.
;      STREET:  1900 Fifteenth Street
;      CITY:  Boulder
;      STATE:  CO
;      COUNTRY:  USA
;      ZIP:  80302
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/651,155B
;      FILING DATE:  17-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Petersen Mr., Steven C.
;      REGISTRATION NUMBER:  36,238
;      REFERENCE/DOCKET NUMBER:  17060.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303/546-1300
;      TELEFAX:  303/449-5426
;      TELEX:  ABA1475
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  438 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-651-155B-134

  Query Match             59.3%;  Score 17.8;  DB 3;  Length 438;
  Best Local Similarity   75.9%;  Pred. No. 43;
  Matches   22;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATGAC 30
              | |||| |||  | ||||||| | |||||
Db        149 ATAATGAATGGGCCGTGTACACCCATGAC 177


RESULT 4
US-09-194-036B-134
; Sequence 134, Application US/09194036B
; Patent No. 6548246
;  GENERAL INFORMATION:
;       APPLICANT: Mahan, Michael J.
;                  Conner, Christopher P.
;                  Hiethoff, Douglas M.
;       TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION
;                           OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST
;                           INFECTION
;       NUMBER OF SEQUENCES: 255
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Morrison & Foerster
;            STREET: 755 Page Mill Road
;            CITY: Mountain View
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/194,036B
;            FILING DATE: 17-No. 6548246-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/US97/08208
;            FILING DATE: 1997-05-16
;            APPLICATION NUMBER: US 08/651,155
;            FILING DATE: 1996-05-17
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Shantanu Basu
;            REGISTRATION NUMBER: 43,318
;            REFERENCE/DOCKET NUMBER: 220002060601
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (650) 813-5995
;            TELEFAX: (650) 494-0792
;  INFORMATION FOR SEQ ID NO: 134:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 438 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       HYPOTHETICAL: NO
```

```
;       ANTI-SENSE: YES
;       ORIGINAL SOURCE:
;            ORGANISM: DNA (other)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 134:
US-09-194-036B-134

  Query Match             59.3%;  Score 17.8;  DB 4;  Length 438;
  Best Local Similarity   75.9%;  Pred. No. 43;
  Matches   22;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATGAC 30
              | |||| ||| | ||||||| | |||||
Db        149 ATAATGAATGGGCCGTGTACACCCATGAC 177


RESULT 5
US-09-457-046B-15
; Sequence 15, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
;  APPLICANT: Croteau, Rodney et al.
;  TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
;  FILE REFERENCE: 53679
;  CURRENT APPLICATION NUMBER: US/09/457,046B
;  CURRENT FILING DATE:  1999-12-07
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 908
;   TYPE: DNA
;   ORGANISM: Taxus cuspidata
US-09-457-046B-15

  Query Match             59.3%;  Score 17.8;  DB 3;  Length 908;
  Best Local Similarity   75.9%;  Pred. No. 48;
  Matches   22;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATGAC 30
              |||  | || ||||||  | ||||||||||
Db        775 ACAGAGAATTTCCGGTTAAGATCTATGAC 803


RESULT 6
US-09-248-796A-7260/c
; Sequence 7260, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7260
;   LENGTH: 198
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-7260

  Query Match             58.0%;  Score 17.4;  DB 4;  Length 198;
  Best Local Similarity   77.8%;  Pred. No. 58;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTAT 27
              ||||||| |||||   ||||||| ||
Db        196 AACAATGAATGTCACTTGTACATGCAT 170


RESULT 7
US-08-605-150A-15/c
; Sequence 15, Application US/08605150A
; Patent No. 6103520
; GENERAL INFORMATION:
;   APPLICANT:  Topfer, Reinhard
;   APPLICANT:  Hausmann, Ludger
;   APPLICANT:  Schell, Jozef
;   TITLE OF INVENTION:  GLYCEROL-3-PHOSPHATE DEHYDROGENASE
;   TITLE OF INVENTION:
;   NUMBER OF SEQUENCES:  17
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Klein & Szekeres
;     STREET:  4199 Campus Drive, Suite 700
;     CITY:  Irvine
;     STATE:  CA
;     COUNTRY:  USA
;     ZIP:  92715
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/605,150A
;     FILING DATE:  01-MAR-1996
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  PCT\EP94\02936
;     FILING DATE:  02-SEP-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  DE P4329827.3
;     FILING DATE:  03-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Szekeres, Gabor L.
;     REGISTRATION NUMBER:  28,675
;     REFERENCE/DOCKET NUMBER:  542-04-PA
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  714-854-5502
;     TELEFAX:  714-854-4897
;  INFORMATION FOR SEQ ID NO:  15:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  1507 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA (genomic)
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   ORIGINAL SOURCE:
;     ORGANISM:  Cuphea lanceolata
;   IMMEDIATE SOURCE:
;     LIBRARY:  Genomic lambda FIX II
;     CLONE:  C1GPDHg9
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  1193..1375
US-08-605-150A-15

  Query Match             58.0%;  Score 17.4;  DB 3;  Length 1507;
  Best Local Similarity   77.8%;  Pred. No. 81;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CAATGTATGTCCGGTGTACATCTATGA 29
              |||||||||||  |||   | ||||||
Db       1018 CAATGTATGTCTCGTGGCAAACTATGA 992


RESULT 8
US-09-221-017B-916/c
; Sequence 916, Application US/09221017B
; Patent No. 6444799
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Ross, Bruce C.
;    TITLE OF INVENTION:  P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
;    NUMBER OF SEQUENCES:  1120
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/221,017B
;      FILING DATE:  23-DEC-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP1182
;      FILING DATE:  31-DEC-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP1546
;      FILING DATE:  30-JAN-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP2911
;      FILING DATE:  09-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU98/01023
;      FILING DATE:  10-DEC-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monroy, Gladys H
;      REGISTRATION NUMBER:  32,430
;      REFERENCE/DOCKET NUMBER:  27340-20021.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-813-5600
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  916:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2182 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  UNKNOWN
;    ORIGINAL SOURCE:
;      ORGANISM:  PORYPHYROMONAS GINGIVALIS
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1...2182
US-09-221-017B-916

  Query Match             58.0%;  Score 17.4;  DB 4;  Length 2182;
  Best Local Similarity   77.8%;  Pred. No. 86;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CAATGTATGTCCGGTGTACATCTATGA 29
              |||||| ||| || ||| ||| ||||
Db       1032 CAATGTGGGTCTGGAGTATATCCATGA 1006


RESULT 9
US-08-750-357-8/c
; Sequence 8, Application US/08750357
; Patent No. 6008437
;  GENERAL INFORMATION:
;    APPLICANT:  KREBBERS, Enno
;    APPLICANT:  WILLIAMS, Mark
;    APPLICANT:  LEEMANS, Jan
;    TITLE OF INVENTION:  USE OF ANTHOCYANIN GENES TO MAINTAIN
;    TITLE OF INVENTION:  MALE STERILE PLANTS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS, L.L.P.
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/750,357
;      FILING DATE:  21-MAR-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McGowan, Malcolm K.
;      REGISTRATION NUMBER:  39,300
;      REFERENCE/DOCKET NUMBER:  018030-100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2662 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-750-357-8

  Query Match             58.0%;  Score 17.4;  DB 3;  Length 2662;
  Best Local Similarity   77.8%;  Pred. No. 89;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              | |||| ||||||| || || |||| |
Db        392 AAAATGAATGTCCGTTGGACTTCTAGG 366


RESULT 10
US-09-112-096-28/c
; Sequence 28, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
;  APPLICANT: Reiner Laus
;  APPLICANT: Michael H. Shapero
;  APPLICANT: Larisa Tsavaler
;  TITLE OF INVENTION: Prostate Tumor Polynucleotide and
;  TITLE OF INVENTION: Antigen Compositions
;  FILE REFERENCE: 7636-0015.30
;  CURRENT APPLICATION NUMBER: US/09/112,096
;  CURRENT FILING DATE: 1998-07-09
;  EARLIER APPLICATION NUMBER: 60/056,110
;  EARLIER FILING DATE: 1997-08-20
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28
;   LENGTH: 3848
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-112-096-28

  Query Match             58.0%;  Score 17.4;  DB 3;  Length 3848;
  Best Local Similarity   77.8%;  Pred. No. 94;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              ||||| ||||| |  | |||||||  ||
Db       2323 ACAATTTATGTTCTTTCTACATCTCTG 2297
```

```
RESULT 11
US-09-112-096-14/c
; Sequence 14, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
;  APPLICANT: Reiner Laus
;  APPLICANT: Michael H. Shapero
;  APPLICANT: Larisa Tsavaler
;  TITLE OF INVENTION: Prostate Tumor Polynucleotide and
;  TITLE OF INVENTION: Antigen Compositions
;  FILE REFERENCE: 7636-0015.30
;  CURRENT APPLICATION NUMBER: US/09/112,096
;  CURRENT FILING DATE: 1998-07-09
;  EARLIER APPLICATION NUMBER: 60/056,110
;  EARLIER FILING DATE: 1997-08-20
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
;   LENGTH: 5668
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-112-096-14

  Query Match             58.0%;  Score 17.4;  DB 3;  Length 5668;
  Best Local Similarity   77.8%;  Pred. No. 1e+02;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              ||||| ||||| |  | ||||||| ||
Db       4143 ACAATTTATGTTCTTTCTACATCTCTG 4117


RESULT 12
US-09-636-215-777/c
; Sequence 777, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 777
;   LENGTH: 5668
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-636-215-777

  Query Match             58.0%;  Score 17.4;  DB 4;  Length 5668;
  Best Local Similarity   77.8%;  Pred. No. 1e+02;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              ||||| ||||| |  | ||||||| ||
Db       4143 ACAATTTATGTTCTTTCTACATCTCTG 4117


RESULT 13
US-09-685-166A-777/c
; Sequence 777, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 777
;   LENGTH: 5668
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-777

  Query Match             58.0%;  Score 17.4;  DB 4;  Length 5668;
  Best Local Similarity   77.8%;  Pred. No. 1e+02;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              ||||| ||||| |  | ||||||| ||
Db       4143 ACAATTTATGTTCTTTCTACATCTCTG 4117


RESULT 14
US-09-679-426-777/c
; Sequence 777, Application US/09679426
; Patent No. 6759515
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
```

```
;  FILE REFERENCE:  210121.427C20
;  CURRENT APPLICATION NUMBER: US/09/679,426
;  CURRENT FILING DATE:  2000-10-02
;  NUMBER OF SEQ ID NOS:  895
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 777
;   LENGTH: 5668
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-679-426-777

  Query Match             58.0%;  Score 17.4;  DB 4;  Length 5668;
  Best Local Similarity   77.8%;  Pred. No. 1e+02;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              ||||| ||||| |  | ||||||| ||
Db       4143 ACAATTTATGTTCTTTCTACATCTCTG 4117


RESULT 15
US-08-920-422-17
; Sequence 17, Application US/08920422A
; Patent No. 6255473
; GENERAL INFORMATION:
;  APPLICANT: Vitek, Michael P.
;  APPLICANT: Mitsuda, No. 6255473iaki
;  APPLICANT: Roses, Allen D.
;  TITLE OF INVENTION: Presenilin-1 Gene Promoter
;  FILE REFERENCE: VITEKPRESENILIN
;  CURRENT APPLICATION NUMBER: US/08/920,422A
;  CURRENT FILING DATE: 1997-08-29
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 48974
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-08-920-422-17

  Query Match             58.0%;  Score 17.4;  DB 3;  Length 48974;
  Best Local Similarity   77.8%;  Pred. No. 1.4e+02;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              | | |||||||||  ||| |||| ||||
Db      31022 ATAGTGTATGTCATGTGGACATGTATG 31048


Search completed: December  6, 2004, 22:52:28
Job time : 4.00716 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 19:02:39 ; Search time 17.7259 Seconds
(without alignments)
9300.077 Million cell updates/sec

Title: US-09-806-197-24
Perfect score: 30
Sequence: 1 aacaatgtatgtccggtgtacatctatgac 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3695051 seqs, 2747533894 residues

Total number of hits satisfying chosen parameters: 7390102

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	19.6	65.3	380	17	US-10-437-963-26247	Sequence 26247, A
2	19.4	64.7	374849	13	US-10-087-192-1627	Sequence 1627, Ap
c 3	19.2	64.0	42772	13	US-10-087-192-1903	Sequence 1903, Ap
c 4	19.2	64.0	2731748	17	US-10-297-465A-1	Sequence 1, Appli
5	19	63.3	512	17	US-10-021-323-6008	Sequence 6008, Ap
6	19	63.3	50000	14	US-10-152-724A-22	Sequence 22, Appl
c 7	19	63.3	1601042	13	US-10-027-632-59064	Sequence 59064, A
c 8	19	63.3	1601042	15	US-10-027-632-59064	Sequence 59064, A
c 9	18.8	62.7	1630	9	US-09-822-830A-444	Sequence 444, App
c 10	18.6	62.0	665	13	US-10-027-632-8033	Sequence 8033, Ap
c 11	18.6	62.0	665	15	US-10-027-632-8033	Sequence 8033, Ap
12	18.6	62.0	2152	15	US-10-372-095-3	Sequence 3, Appli
13	18.6	62.0	90442	14	US-10-105-637-1	Sequence 1, Appli
14	18.6	62.0	90442	15	US-10-034-650-43	Sequence 43, Appl
c 15	18.4	61.3	255	16	US-10-424-599-132779	Sequence 132779,
c 16	18.4	61.3	650	13	US-10-027-632-320106	Sequence 320106,
c 17	18.4	61.3	650	15	US-10-027-632-320106	Sequence 320106,
c 18	18.4	61.3	662	13	US-10-027-632-84202	Sequence 84202, A
c 19	18.4	61.3	662	15	US-10-027-632-84202	Sequence 84202, A
c 20	18.4	61.3	188017	13	US-10-087-192-1951	Sequence 1951, Ap
c 21	18.2	60.7	40050	16	US-10-052-482-91	Sequence 91, Appl
22	18.2	60.7	310122	18	US-10-417-375-1	Sequence 1, Appli
23	18	60.0	287	16	US-10-621-901-1833	Sequence 1833, Ap
24	18	60.0	287	16	US-10-621-901-1859	Sequence 1859, Ap
25	18	60.0	386	16	US-10-621-901-1927	Sequence 1927, Ap
c 26	18	60.0	401	9	US-09-795-668-1084	Sequence 1084, Ap
c 27	18	60.0	401	9	US-09-795-686-1084	Sequence 1084, Ap
c 28	18	60.0	401	9	US-09-946-807-1084	Sequence 1084, Ap
29	18	60.0	448	16	US-10-621-901-1276	Sequence 1276, Ap
c 30	18	60.0	620	18	US-10-425-115-32472	Sequence 32472, A
c 31	18	60.0	649	16	US-10-424-599-103647	Sequence 103647,
c 32	18	60.0	2194	17	US-10-437-963-21442	Sequence 21442, A
33	18	60.0	2880	16	US-10-424-599-74550	Sequence 74550, A
c 34	18	60.0	42007	15	US-10-085-117-337	Sequence 337, App
c 35	18	60.0	1503841	9	US-09-795-668-1	Sequence 1, Appli
c 36	18	60.0	1503841	9	US-09-795-686-1	Sequence 1, Appli
c 37	18	60.0	1503841	9	US-09-946-807-1	Sequence 1, Appli
38	17.8	59.3	141	13	US-10-033-528-1890	Sequence 1890, Ap
39	17.8	59.3	141	15	US-10-099-926-1890	Sequence 1890, Ap
c 40	17.8	59.3	421	17	US-10-437-963-89493	Sequence 89493, A
c 41	17.8	59.3	428	14	US-10-198-846-922	Sequence 922, App
42	17.8	59.3	500	10	US-09-918-995-23954	Sequence 23954, A
c 43	17.8	59.3	662	14	US-10-198-846-829	Sequence 829, App
c 44	17.8	59.3	881	14	US-10-198-846-10611	Sequence 10611, A
45	17.8	59.3	908	9	US-09-866-572A-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-10-437-963-26247/c
; Sequence 26247, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 26247
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(380)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_31056C.1
US-10-437-963-26247

Query Match 65.3%; Score 19.6; DB 17; Length 380;
Best Local Similarity 78.6%; Pred. No. 48;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 AACAATGTATGTCCGGTGTACATCTATG 28
              || |||| || |   ||||||||||||
Db         54 AAAAATGCATCTNANGTGTACATCTATG 27
```

RESULT 2
US-10-087-192-1627
; Sequence 1627, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1627
; LENGTH: 374849
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(374849)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1627

Query Match 64.7%; Score 19.4; DB 13; Length 374849;
Best Local Similarity 79.3%; Pred. No. 1.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 AACAATGTATGTCCGGTGTACATCTATGA 29
              |||||||||||| | ||| |||  ||| |
Db       6156 AACAATGTATGTTCAGTGAACAGGTATTA 6184
```

RESULT 3
US-10-087-192-1903/c
; Sequence 1903, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1903
; LENGTH: 42772
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(42772)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1903

Query Match 64.0%; Score 19.2; DB 13; Length 42772;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          5 ATGTATGTCCGGTGTACATCTATG 28
              ||||||||  |||||||||  ||||
Db      15706 ATGTATGTGTGGTGTACATGTATG 15683
```

RESULT 4
US-10-297-465A-1/c
; Sequence 1, Application US/10297465A
; Publication No. US20040142413A1
; GENERAL INFORMATION:
; APPLICANT: Simpson, Andrew
; APPLICANT: Reinach, Fernando
; APPLICANT: Setubal, Joao
; APPLICANT: Medianis, Joao
; APPLICANT: Arruda, Paulo
; TITLE OF INVENTION: Isolated Genome of Xylella Fastidiosa and Uses Thereof
; FILE REFERENCE: FAPESP 202 US (10213376)
; CURRENT APPLICATION NUMBER: US/10/297,465A
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01618
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,906
; PRIOR FILING DATE: 2001-06-17
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2731748
; TYPE: DNA
; ORGANISM: Xylella fastidiosa
US-10-297-465A-1

Query Match 64.0%; Score 19.2; DB 17; Length 2731748;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 AACAATGTATGTCCGGTGTACATC 24
              ||||||  |||||||  ||||||||
Db    1072833 AACAATAAATGTCCGCTGTACATC 1072810
```

RESULT 5
US-10-021-323-6008
; Sequence 6008, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 6008
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-012-Q1-N6-C3
US-10-021-323-6008

Query Match 63.3%; Score 19; DB 17; Length 512;
Best Local Similarity 81.5%; Pred. No. 96;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          3 CAATGTATGTCCGGTGTACATCTATGA 29
              ||||  || ||  ||||||||||| |||
Db        394 CAATCCATTTCTGGTGTACATCTCTGA 420
```

```
RESULT 6
US-10-152-724A-22
; Sequence 22, Application US/10152724A
; Publication No. US20030082714A1
; GENERAL INFORMATION:
;  APPLICANT: LITTLE, Melissa
;  APPLICANT:  HOLMES, Gregory
;  APPLICANT:  KOLLE, Gabriel
;  APPLICANT:  YAMADA, Toshiya
;  APPLICANT:  GEORGAS, Kylie
;  APPLICANT:  WILKINSON, Lorine
;  TITLE OF INVENTION: No. US20030082714A1el Nucleic Acid and Polypeptide
;  FILE REFERENCE: P22378
;  CURRENT APPLICATION NUMBER: US/10/152,724A
;  CURRENT FILING DATE:  2002-05-23
;  PRIOR APPLICATION NUMBER: Australian App No. US20030082714A1PQ 4348
;  PRIOR FILING DATE: 1999-11-26
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 50000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-152-724A-22

  Query Match             63.3%;  Score 19;  DB 14;  Length 50000;
  Best Local Similarity   81.5%;  Pred. No. 1.9e+02;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 AATGTATGTCCGGTGTACATCTATGAC 30
              ||||| |||  | |||||||||| ||||
Db      33196 AATGTGTGTTAGATGTACATCTGTGAC 33222


RESULT 7
US-10-027-632-59064/c
; Sequence 59064, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59064
;   LENGTH: 1601042
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1601042)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-59064

  Query Match             63.3%;  Score 19;  DB 13;  Length 1601042;
  Best Local Similarity   81.5%;  Pred. No. 3.1e+02;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              ||||||| ||| ||||  | |||||||
Db     707259 ACAATGTGTGTGCGGTCAAAATCTATG 707233


RESULT 8
US-10-027-632-59064/c
; Sequence 59064, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59064
;   LENGTH: 1601042
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1601042)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-59064

  Query Match             63.3%;  Score 19;  DB 15;  Length 1601042;
  Best Local Similarity   81.5%;  Pred. No. 3.1e+02;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              ||||||| ||| ||||  | |||||||
Db     707259 ACAATGTGTGTGCGGTCAAAATCTATG 707233


RESULT 9
US-09-822-830A-444/c
; Sequence 444, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
;  APPLICANT: Genetics Institute, Inc.
;  APPLICANT:  Wong, Gordon G.
;  APPLICANT:  Clark, Hilary
;  APPLICANT:  Fechtel, Kim
;  APPLICANT:  Agostino, Michael J.
;  APPLICANT:  Howes, Steven H.
;  APPLICANT:  Resnick, Richard J.
;  APPLICANT:  Gulukota, Kamalakar
;  APPLICANT:  Graham, James R.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
;  FILE REFERENCE: GIN 6402
;  CURRENT APPLICATION NUMBER: US/09/822,830A
;  CURRENT FILING DATE:  2001-03-29
```

```
;  PRIOR APPLICATION NUMBER: 60/195,604
;  PRIOR FILING DATE: 2000-04-06
;  NUMBER OF SEQ ID NOS: 631
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 444
;   LENGTH: 1630
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 535
;   OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-444

  Query Match             62.7%;  Score 18.8;  DB 9;  Length 1630;
  Best Local Similarity   76.7%;  Pred. No. 1.4e+02;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AACAATGTATGTCCGGTGTACATCTATGAC 30
              ||| | ||| | || |||  ||||||||||
Db        633 AACGAAGTAAGCCCTGTGATCATCTATGAC 604


RESULT 10
US-10-027-632-8033/c
; Sequence 8033, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8033
;   LENGTH: 665
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-8033

  Query Match             62.0%;  Score 18.6;  DB 13;  Length 665;
  Best Local Similarity   77.8%;  Pred. No. 1.5e+02;
  Matches   21;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              |||||||||||  |  :|| ||| ||||
Db        459 ACAATGTATGTGCTAYGTGCATGTATG 433


RESULT 11
US-10-027-632-8033/c
; Sequence 8033, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8033
;   LENGTH: 665
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-8033

  Query Match             62.0%;  Score 18.6;  DB 15;  Length 665;
  Best Local Similarity   77.8%;  Pred. No. 1.5e+02;
  Matches   21;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATG 28
              |||||||||||  |  :|| ||| ||||
Db        459 ACAATGTATGTGCTAYGTGCATGTATG 433


RESULT 12
US-10-372-095-3
; Sequence 3, Application US/10372095
; Publication No. US20030162256A1
; GENERAL INFORMATION:
;  APPLICANT: Juppner, Harald
;  APPLICANT:  Rubin, David A.
;  TITLE OF INVENTION: PTH1R and PTH3R Receptors, Methods and Uses Thereof
;  FILE REFERENCE: 0609.4740002
;  CURRENT APPLICATION NUMBER: US/10/372,095
;  CURRENT FILING DATE:  2003-02-25
;  PRIOR APPLICATION NUMBER: 09/449,632
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: US 60/110,467
;  PRIOR FILING DATE: 1998-11-30
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 2152
;   TYPE: DNA
;   ORGANISM: Danio rerio
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (394)..(2019)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2125)..(2125)
;   OTHER INFORMATION: n is any nucleotide of a,t,g or c
US-10-372-095-3

  Query Match             62.0%;  Score 18.6;  DB 15;  Length 2152;
  Best Local Similarity   84.0%;  Pred. No. 1.8e+02;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 TGTATGTCCGGTGTACATCTATGAC 30
              | | ||||| | ||||||||||||||
Db        639 TCTGTGTCCTGAGTACATCTATGAC 663
```

```
RESULT 13
US-10-105-637-1
; Sequence 1, Application US/10105637
; Publication No. US20030087252A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David
;  APPLICANT:  Engelhard, Eric
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER
;  TITLE OF INVENTION:  ASSOCIATED WITH ALTERED EXPRESSION OF PRDM11
;  FILE REFERENCE: 529452500120
;  CURRENT APPLICATION NUMBER: US/10/105,637
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 10/034,650
;  PRIOR FILING DATE: 2001-12-20
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 90442
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-105-637-1

  Query Match             62.0%;  Score 18.6;  DB 14;  Length 90442;
  Best Local Similarity   84.0%;  Pred. No. 3.2e+02;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 ATGTATGTCCGGTGTACATCTATGA 29
              ||||||||| ||| | |||| |||||
Db       3731 ATGTATGTGCGGGGCACATATATGA 3755


RESULT 14
US-10-034-650-43
; Sequence 43, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David
;  APPLICANT:  Engelhard, Eric
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000128
;  CURRENT APPLICATION NUMBER: US/10/034,650
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 09/474,377
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43
;   LENGTH: 90442
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-034-650-43

  Query Match             62.0%;  Score 18.6;  DB 15;  Length 90442;
  Best Local Similarity   84.0%;  Pred. No. 3.2e+02;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 ATGTATGTCCGGTGTACATCTATGA 29
              ||||||||| ||| | |||| |||||
Db       3731 ATGTATGTGCGGGGCACATATATGA 3755


RESULT 15
US-10-424-599-132779/c
; Sequence 132779, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132779
;   LENGTH: 255
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_90906C.1
US-10-424-599-132779

  Query Match             61.3%;  Score 18.4;  DB 16;  Length 255;
  Best Local Similarity   78.6%;  Pred. No. 1.7e+02;
  Matches   22;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ACAATGTATGTCCGGTGTACATCTATGA 29
              || |||| |||||| || ||| | ||||
Db        105 ACTATGTTTGTCCGTTGGACAGCGATGA 78


Search completed: December  7, 2004, 06:11:27
Job time : 23.7259 secs
```